

KTAIRCNBPREKNGKXVCGRMKFKSNTPEPLCKOKRDFREDOCAHNRDGHKFNING
 LDPRKRWPKISGILMKDKCKLPEVAGNTAYQLDRIDTGPCGQIDPNTDCLDGLC
 RQAGCDHVLNLSKARNDKCGVCGGDNSSCKTQVAGTFTVYGNVTVVRIADGATNIDL
 QHSFGEEDDDVYLAISSKGFELNNGNVYMAKKEIIGAVVVEYSSETFAVERIN
 SINDRIEELLQVLYGKLYNPDVRSFNIPIEDKOCQYOMNSHGMWQCSKPKCOEER
 KKIIVCTRESDDLVSDORCDLPOBGIITFCGDDCCDLRMHVAVSSSECNACCGGGR
 TLDIYCAKYSRDGKTEKVDGFCSSHPKPSREKSGEENCGRRYSMAWTECSKSKD
 GGTORRRAICVTRNDVLDLDDSKTNOEKYTLTORGSEFPQMSGDMSCDLVTCGKH
 KIRQVWCGQEDRLNDRMCQDPRTKPSMOTQOPEASQWAGPWTGQSVTCGQYQLR
 AVKCIIGTYSVVDNDNCAARPTDODCELPSCHPRPAABETRRSTYSAPRTQWR
 GSWTCSATCGGTRMYKVSCHRDENGVSADESAATLPPRVKESVYPCQGMKLD
 RSCSVTSGGRATRVQVAVNSDHYIDRESDODYIPEITDODSINSRPPORPDDGL
 AOHPPONEDYRBRASPSRTHYLIGGNMRTGPMWAGCSTPAGLSORRYVQOCDENGT
 ANDCVERTKRDQKACSESPRCQWATGNNBECTKICGSGTTRLYVYQCSNNGRERFDL
 SCEILDKPRDQCNTHACRNDAAWSTGWSSCSVSCGGRNQRVNYCMAKKSSEHLS
 DYCCKILAKNHRKRCGRRCRKAAGWQVCGVCGRQVQOHRVQGLDITNHTARTLE
 CNPVTREPEREDCGPRCLYTWRAEMEQECRTEGSEGRSRKRVVGVDDNKKNEVHGAR
 CDSKRPVYRECSLOPCSEYVITGEMSVVPMEL"

BASE COUNT 1471 a 1112 c 1363 g 1194 t
 ORIGIN
 AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803
 Final Score = 7 Optimized score = 14 Significance = -0.35
 Residue Identity = 21% Matches = 20 Mismatches = 46
 Gaps = 29 Conservative Substitutions = 0
 Translation Frame = 6

IGRVASSLPHSCNHPSTNPSATLREWLMSXKRWLDFRRGXYLXRFSCDRDILAGLQTTAFTVHRPLNHTVACRPMH
 210 230 230 240 250 260 270 280
 LVFVYVHTNHIAVGAFAAGIGAFRLPLCPREVSEGTWALAVAFRLRSQWVWALGLSGYFCSLDLTAHMRPL
 290 300 310 320 330 340 350
 YASATGHRALTPRASFPFGAASSSALSVPMMLSQVLTVTFYTXMASIFCHAVNISLFPSTTRDRTRARAST
 360 370 380 390 400 410 420
 PCSYVWMSVCVTLDTIGRFLKFNFAOIKWPRTYGRPLRDNQSCSYASTTQLSALSPVAISPLRTRAGTGGLL
 430 440 450 460 470 480 490
 IRRVSLHTVVCGVSIFFLYNNPRLPGSASTGTGTGSPGASSPLVSTEMGMAAGAABGTVLILLEGVLSXAT
 500 510 520 530 540 550 560
 WMGRLRAWXXHGTVLVSWMDIILVLPIDHVVYVVDVHNHLPGCCPTLTLGHRARARAVQGIPLPTGCHHRTFFLC
 570 580 590 600 610 620 630 640
 HWSBQSGTGLVSHRALLISAADVSHRGTPRTSGXAWGRPRKSPFLGSMCTVYCASPRFRGSKWRMTXXYFTVL
 650 660 670 680 690 700 710
 GISWSSCCIVIVYIYKHHSIPNDAFHCSXLYSLSTSPFALSPTGRILPCCFTFRLLTSHRRCMLGIRVNHSTIX
 720 730 740 750 760 770 780
 SIPTKRLPDLAVLMPFSTGDAALXVPSRFRPLMTRELTARPLNGLFLMCARAVLDYIISGIDTNSPSPLGRPV
 790 800 810 820 830 840 850
 TAFKTFSPGRIARATRTVFPFXAFPTVAMFGMAAKTINILSLPDRALGTGVVQCAVWQATLIGTFTTFRAG
 860 870 880 890 900 910 920
 NMPROVTVCTTGTGSNVSRLGQOPASLIRNSKLIIRFGANKRSPFPPLAGFTACLPWEMTVVFKLLRFTFMWN
 930 940 950 960 970 980 990 1000
 IERTAVYGVVQLSHRQWLKQKFLMLNACXVNSFGSLGFPVYLHNSIPNANPFGCHCNRKKSIXOEFPTTAXX
 1010 1020 1030 1040 1050 1060 1070
 SXVYVIVCFPEKTEVLPKPHINIGSTWSNDSHDSIVITMVCIKSCCHFAKRIIATVPTFTISPGFXVXNNIASSL
 1080 1090 1100 1110 1120 1130 1140
 PAKALDIDICVLATRRSSIVHSKLLIVGCVPCHSABOLAPVLHONSTVFREDPAHIGKQFVDMKMLRVLVXVST
 1150 1160 1170 1180 1190 1200 1210

X
 QLP-----
 11
 1200

10 20 30 40 50 60
 GILYVWKKOE-----LTFGREGVQVPRYMDORLIMCNV---NGLVHKGCRTGHPWMDGTECEFGKNC-
 LEFSEVYILLEETWLRVAGLGFHSTSPS---YTFVSTIFWFMVSVALNSNCFDAPSTCSGASGSKGPTSPGSIQH
 1220 1230 1240 1250 1260 1270 1280
 -----KYGCV
 X
 GDHIFGNGTFSILAVLSRLRLRPIRGPCVSLPAPAFVYSIDVIADEPESLHHIWHNLRTWSKNQPTLREIHVVK
 1290 1300 1310 1320 1330 1340 1350
 DAWMLDRORYGSSGRFRFQKTLATISVYKSLSDIPSTLTXPRPGVVEVQCSMDLGTLLNSRFFVAVVIMRH
 1360 1370 1380 1390 1400 1410 1420
 VKHMAOLMGDRKSCSTOSTIFTRNPAISIGTNGSOPSOAXGITPVSASADLLSCXENSSIMWDSITWTVLRMLP
 1430 1440 1450 1460 1470 1480 1490 1500
 LAKYFGCCLSTIKRYGRPLRFLMHNHNVHNNININISMVWFVIDBGVYKXKXNTVLKVFPSMVVXNSVYVCHHQ
 1510 1520 1530 1540 1550 1560 1570
 DFKYKRWIGXKTFPCSPVGRPLFSCVYVRLITISRKLCCXAAVYXCVYTSQVNPFPSPSCSGFLVFTVPIFWFX
 1580 1590 1600 1610 1620 1630 1640
 GVTGMPSCXGSLISGGAPRINDVYGFVLFLFLFTLHRLXWPFKIIPIIRPBECAOHSAEADDRYVIG
 1650 1660 1670 1680 1690 1700 1710

3. US-09-972-467-2 (509-578)
 ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
 TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
 DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
 ACCESSION AB037733
 VERSION AB037733.1 GI:7242978
 KEYWORDS
 SOURCE Homo sapiens brain cDNA to mRNA, clone_11b:pbuescript11 SK plus
 clone_fh11767.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (sites) Nagase,T., Kikuno,R., Ishikawa,K.T., Hirosewa,M. and Ohara,O.
 Prediction of the coding sequences of unidentified human genes.
 XVI. The complete sequences of 150 new cDNA clones from brain which
 code for large proteins in vitro
 DNA Res. 7 (11), 65-73 (2000)
 JOURNAL MEDLINE 20181126
 REFERENCE 2 (bases 1 to 5139)
 AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
 TITLE Direct Submission
 JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba
 292-0812, Japan (E-mail:cdna@infokazusa.or.jp,
 URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
 Fax:+81-438-52-3914)
 FEATUES
 location/Qualifiers
 1..5139
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="fh11767"
 /tissue_type="brain"
 /clone_lib="pbuescript11 SK plus"
 1..4417
 /gene="KIAA1312"
 <1..4417
 /gene="KIAA1312"
 /note="Start codon is not identified."
 /codon_start=2
 /product="KIAA1312 protein"
 /protein_id="BAA92550.1"


```

/codon_start=2
/product="K1A1312.protein"
/protein_id="BA92550.1"
/db_xref="GI:7242979"
/translation="NSEHVAVLSLCSGMLGTRFSDGDYETLEPLOSMDQEDQEDQBNK
PHIYRRSAPQREPTGHRHACTSEHKRNSKDKKTKARKKGEINLACGVALNSG
LAEAFSAVGNKTDNFRKRRRTRKFLSYRFEVLYRVAEYVAVLWVHGHMDFIIL
TKMSIAYIKDPSISGLNLIYVNLVYHNSDDPSTISENMQTLTKRNEQMOHSKNS
PGGIHDDTAVLLTRDODIGRHKDCTLGAELTICDPRRSLTSEEDSGLSAFLTAH
ELGHVYNNMHRDNNKCKEYKSRQHVAPATLNPYTNPMNSCKSRKLTTEPDIATG
ECLLNDRSRKPLVQLGKILYNNKQKELTFGFSQVCPYPMCKRMLKMNLYNVGH
KGRQTRHPADDTCEBPGKHCYKFCVPEKMDVLTQBSWMSWPCFSTKCGGI
KTAIREGNRPEKNGKCYGGRMKKSCNTPRCLKQVDFVDFECAFDFKHPFING
LDPNVWVPEYSGILMKDRCKLEFVAGNTAYLQDRYIDDTPCGDPTNDCVQGLC
ROAGCDHVLNMSKARBDKCGVGDGNSCKTVAGTNYVYVYVYVYVYVYVYVYVYVYV
HSPSEGTDDDDVYALSSSGPELLNGVNYVYVYVYVYVYVYVYVYVYVYVYVYVYV
STPRIQEBLLDYLTVGKLYNPDYKRSFNIPEDKPODFYNNHSPWQACSKPCQDGR
KRLVLCRESDQLVSDQKDRLPQGHITTEPGDCLRMHVARSBSAQKGLYR
TLDYGAQKSRDLGKTEKVDGFCSSHPKPSRREKCSGECNTGMRYSAMTECSKSD
GTQRRRAICVNRNDVLDLSDCTQHEKVTIORSEFPQPKWSDMSWPCFSTKCGGI
KHRQVWCFQEDRLNDMCDDEPKTPTSMQTCQAPCAQWQAPMGQCVTCGQQLR
AVKCIIGYVWYVVDNDCAATVREPDTQCELPSCHPPEPAAEFTSRSTYSAAPDQWR
GSMTPGSAFGCGKTRMRYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV
GSMTPGSAFGCGKTRMRYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV
WSSCVTCGQGRATKQVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV
AQPFDQEDRPRASPSRHYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV
ANDCVBRIRKPEDEGRACEGPPQWYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV
SCBILDKPRDEQCNTHACPHDAAMSTGMSGCVSVCGRGHRKQVYVYVYVYVYVYVYVYVYV
DYCKHLAKPHGHRKCRGRCPKYKAGAMSGCVSVCGRGHRKQVYVYVYVYVYVYVYVYVYV
CNDYTRPESRDCQGRCPDLYTMRABEMOECTKCTGEGSRYRKYVYVYVYVYVYVYVYVYVYV
CDVSKRPVRESLSLQPCBYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV
BASE COUNT 1471 a 1112 c 1362 g 1194 t
ORIGIN

```

```

AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803
Initial score = 6 Optimized score = 13 Significance = -0.38
Residue Identity = 20% Matches = 15 Mismatches = 55
Gaps = 4 Conservative Substitutions = 0
Translation Frame= 5

```

```

XLLLVYHKKGRALCTGRYMKWLYRYFGYRKLFFSSSEAVCXGTGTHLVHERGLFADYVNLKPKISAPI
10 20 30 40 50 60 70
FKLGNLTYXSRPLASLEKSGLVTLDDTISAMPQLKXRRRCSLHTFACASLSDQLPSLLVYSLALQYFILLNCSQA
80 90 100 110 120 130 140
PHSILXVCSPCNNEFFMPIGHQGSQNVYDKGLRDHLHLSRFAKYKPEFLYKGAEDLESNTQFTVYVANSGLK
150 160 170 180 190 200 210
SPTFLHSSLSHSLCHSQAMAEITWYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV
220 230 240 250 260 270 280
GQVCPYPMQCRRLKCNVNVYKGCGR--TQHTPWADGT--ECPERGHCKKCYKCYV 70
HTRPCGTMSTRSRMSCLLAIPLPSKCRQDVTGLDSRVPTRPWCMSGCTRSLSMLRCVYFSDSPHASVYRLGHT
290 300 310 320 330 340 X 350 360
OSTDSKRLDLSIWIGLITGTFCAHVAAHPGAVSNHFLNGFHLPLPSCNHFVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV
370 380 390 400 410 420 430
KAVCYTADHREYVQEFHSSNLETVYHRWADRRVYLFILCIIHHTAAXCTLPSCHKKPTEDKRIHRLFAHQVLYFSP
440 450 460 470 480 490 500
HSRLKIHHPDIQHAAGIRQHRYYMNLPRGQFSTGFHRAHGCGWGRSGDSDPHFGKALSHCLGSPFEG
510 520 530 540 550 560 570 580
AVTWNSPQGFGLYVNFQHTNSDRSRGCHSCXHTSLAGLRYLAHRSQSSKSSPSPSIAHHRVSONIILPLPLVYXGX
590 600 610 620 630 640

```

```

HRHNSQPSHSHLSSXRISSGYSLHKMLSMGSKTQIATGFLVHOMCSPSSPQGLGDDMMVTHSPGYQLVYL
650 660 670 680 690 700 710 720
LHYSCHLPLRFLYKSOXCSISLLAGIRVYKXHCYVPRDPRARMMHIRPAADKSAKMLAMSGHTFTHLIDLRQTD
730 740 750 760 770 780 790
TRPGAVYALFRRXSTLTLISLQISYVDKGTCTSEWXLISLSDVLCICPHVHNFEGYXHKRKPSSGSHRHSFLNIQ
800 810 820 830 840 850 860
SRONSASHPYIPLSTIFPGLVMDGCGCKNHRQSPSSHPACTYTHNSRCPMCSGNSPTGHYHNSQOHALTSGH
870 880 890 900 910 920 930
SLYHFVQXVOAGAADRIFDQKQXADQIPMCKQVYFVPLRVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV
940 950 960 970 980 990 1000
RCSTTPRPFKREAVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV
1010 1020 1030 1040 1050 1060 1070 1080
LPLRLNCAALHQYWHQIEFEGPQYCNHNVLYXMFLEPFCMKNYRHHKPHIYLSGLLSLKHDRITQLAGKGRGH
1090 1100 1110 1120 1130 1140 1150
RYHLCPGKHFHLSLCEADSRILCSLPLCRTTCTGPPSEPHCTIXGSPAHWEADRCXVNSROSEHTVHLGSL
1160 1170 1180 1190 1200 1210 1220
ASAXDAPCCRTXYSFVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV
1230 1240 1250 1260 1270 1280 1290
HTCSAFAQRFPSHREPRWCAEFGSLCVLHXYRCTRAVCTASTAGYTPEN
1300 1310 1320 1330 1340

```

```

5. US-09-972-467-2 (509-578)
ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
TOIG of: ab037733 check: 4803 from: 1 to: 5139

```

```

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
ACCESSION AB037733
VERSION AB037733.1 GI:7242978
KEYWORDS
SOURCE
ORGANISM Homo sapiens
clone:fh11767.
Homo sapiens brain cDNA to mRNA, clone_11b:pbluescriptII SK plus
clone:fh11767.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (sites)
Nagase,T., Kikuno,R., Ishikawa,K.I., Hirozawa,M. and Ohara,O.
Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
DNA Res. 7 (1), 65-73 (2000)
JOURNAL
MEDLINE
20181126
REFERENCE
2 (bases 1 to 5139)
Ohara,O., Nagase,T. and Kikuno,R.
Direct Submission
Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
297-0812, Japan (E-mail:cdna@info@kazusa.or.jp.
URL:htp://www.kazusa.or.jp/dnuge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
Location/Qualifiers
FEATURES
source
1..5139
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="fh11767"
/tissue_type="brain"
/clone_lib="pbluescriptII SK plus"
1..4417
/gene="KIAA1312"
<1..4417
/gene="K1A1312"
/note="Start codon is not identified."

```

```

/codon_start=2
/product="K1AA1312 protein"
/protein_id="BA092550.1"
/db_xref="gi:7242979"
/translation="NSEHTAVISLCSGMLGTFRSHDGDYFTEPLQSMDEDEEONK
PHIYRARSAPQREPTSGHACDTSEHKNRHSKDRKTRARRKMGRIINAGDVALNSG
LATEFASVGNKTDNTPREKRVNRRKRELYPRVEVLVAVDNTMVMVYHGHNLOHYL
TMSIVASTYKDPSTGNINIVYLVNLIYHNEODGPTISENAOITLTKNFCOMOSKRS
PQGHNHDVAVLTLTKDDICRANHCDDTGLABELGTCDFRYSGSISESGSLSTATIH
ELGVFNMRHNDNNKCKEYKSPQHWAPLPIFYGVQVMSKCSKRYITFELIDVIG
ECLLNEBPSRPLVPLVPLGILYVNNKCELLIFGYSQVYMMQCKRMLNWNVNG
KGCQHTPMADGTECEBPKHCKYGVKEMADYDVLDSGMSYSPGTCSTRTGGGI
KTAIECRNRPKNGKGYKVGRRKMKFKCNTPEKQKRDREDDQCAHFDKHNING
LLPNVRYWYKSGIIMDKRCKLFCRVAGNTPAYLDRVLDGTPGDDINDICVOGIC
RQAGCDHYLNKARADKCGVGGSSCKYVAGTPEYVYVAVLVPYIPGATINDVR
OHSEGETDDVDYLAISSKGEFLNGLVYVYAKREIRINAYAYEYSGSEPTAYERN
STRLEQELLDVLSVGLYNDVAKYSINLITEDPQDFINNSHGWACSKPQDGER
KRKLVCTRESQDLVSDQCDRLPQPHITPECGTDCDLRWHVNSRSESKQCGLGR
LTDIYCAKYSRLDGTKEKYVDGFCSSHPKPSNRSCECMTGWRYSAMWTECSGD
GGTORRAICVMTNMDVLDSDSKCTHOEKVYTORCSEFPQMSGSDSVLTCGKCH
KHROYVCOFGEEDRLDMCDEFTKPTSMOTQOPEFCASMOAGPMSGCVTCGOGYOLR
AYKCTIGYMSVYDNDNCAATRRPTDODCELPSGHPRPADETRRSTYSAPRTQWR
GWTFCSATCGKGTMRKYVSCDENGSYADDSACTLPRYAKRECSYVTPCGOMKALD
WSSCSVTGQGGKATQVWCVNSDHYIDRSBEDDYIETEDQDSMSFCPQRTDSEGL
AQHPQNEDEYRPRRSASPSRTHYLGNQMTGPMWACSSSTGAGGQSRVYVQDIBNGTL
ANDCVERIKPDEQACESGPCPOMAVYGMWGECTKICGGGITRILVYCCORRNGEPRDL
SCETLIDKPRDEQCNTHACPHDAAMSTGPMSSCSYSCGRGKORRVYCMADGSHLES
DYCKHLAKPHGRKCGRGRPKWAKAGVSCSVSGRGOORHVGCCOIGTCKIKARETE
CNPYTRPSEPCQSPRRPLIYTRAKEMQDEECTKCTGSEBSRTRKYVVCVNDNKNVEVHGAR
CDVSKRPVDRSECSLQPEYVWITGEMSEVPSWEL"
BASE COUNT      1471 a 1112 c 1362 g 1194 t
ORIGIN

```

```

AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ...
Initial Score = 6 Optimized Score = 12 Significance = -0.38
Residue Identity = 17% Matches = 16 Mismatches = 34
Gaps = 20 Conservative Substitutions = 0
Translation Frame= 1

```

```

QLRNHNQRLRLRAGHILRVYXWGLFYXHTTVYGYTRAKRGTKQTRHHLXQAPRPERLNRKAKMHLRPTOK
10 20 30 40 50 60 70
80 90 100 110 X 120 130 140
XAOXROBENOSKMKGRKDXGWMXSSITKORLNSNNGIFCLMXXDGOHKKRKEDEPKDKTFEFLTLSTVCCSLGGRR
40 50 60 70
-----RLMGN-NVNGVHKGCRTQHTPWADGTECE-----PGKHCKYGFVY
OQNGFIYRKRKSTLYFNPVNNCSLYLKRPRKYWKNNXYCYCELNDSXXGTWAFHIFKCSNRIKLLRVAFAE
150 160 170 180 190 200 X 210
BQSRWNPDSYXCCSLNKTGYLSSHQMKYLLRPGXTGNHLXSLKLFYXXKXYIEYSEYVDRPACVYXNHSXX
220 230 240 250 260 270 280
OQOXXRRRSXSPRACHGNSNFEHLHQPLDLYVYKYSKIKYHNVYRHMVLRVFAXXPTXIQTLRPFACPTARRHPLQRE
290 300 310 320 330 340 350 360
XTMXIDFWTRFSGVPIYDAVQFALVQXKRWSTORLPDSANHTLGRHWDGVRAMKALQWITLCSQRNGCPRDRWI
370 380 390 400 410 420 430
LGLLESIMNLLQNMWRGHNHSHRVOOTRRFKKWKILCPRFXNEILXVLYOHGMSQAEARLBRXTVYCSLXRPAP
440 450 460 470 480 490 500
XHOASASQCALGRPIYIQWNSDEGRVQVYVYLOSREHSLLSASRQSDRWMSLMPGRKXYYLCPGRLDPA5MWRSGCFK
510 520 530 540 550 560 570
LKSPEKXMMGLMWRXFFEQNSGRNITXXSTLMLQYCGPNS5WCYQXCAAOEFLRGNRRXDL5F IKOXRXL
580 590 600 610 620 630 640

```

```

AKMKLCSNGQKGN5HECCSGRVQWVBCRRKXNHLNRSXARTFASGFGVGGKVVQRCFLLEFQYSMKXRTISA
650 660 670 680 690 700 710 720
VLLQSNAMASMOXTPDRGFEETKCLHOGITXSNVCFYSKMKRSAAAPMTHTYXJLMLYRLXPEVACQOQEXMKXP
730 740 750 760 770 780 790
VWLGIPRNLGHLICQIXQAGWEDXEGXWVFLQDPSQTKOPKXMKLGMXHGWLALFCLDXMKFKLIXRWDPEREK
800 810 820 830 840 850 860
YLCQYRKKCTGXOQMHNSRESYHSEVQVPLSTVYETIRLVYLVGHLMLKRAKXAPRGLV5VWXR5IKXXNVPX
870 880 890 900 910 920 930
SSP5CGNEEKHICQTKNPAVIAWVLDPMLSHLWERYPDEIQDLPRXEMLCSGRSCLCPAXXTSGKRMFCDT
1010 1020 1030 1040 1050 1060 1070 1080
LWAMEGLLELLDLDLWARKXGNPNSDVCQDQXPRDRSEYXHPGLYPRNXYGGLFVHTMPSKDPQWML5SAPLP
1090 1100 1110 1120 1130 1140 1150
KXGL5SPRPOPDPHPCARMKRPEVNWPLG5MFFOYLQWRIPRPAACCCSGXKWT
1160 1170 1180 1190 1200

```

```

6. US-09-972-467-2 (509-578)
ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
TOIG of: ab037733 check: 4803 from: 1 to: 5139
LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for K1AA1312 protein, partial cds.
ACCESSION AB037733.1 GI:7242978
VERSION AB037733.1 GI:7242978
KEYWORDS
SOURCE Homo sapiens brain cDNA to mRNA, clone_11b:pBluescriptII SK plus
clone:fh11767.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
Nagase,T., Kikuno,R., Ishikawa,K.I., Hirosewa,M. and Ohara,O.
Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences in vitro
code for large proteins in vitro
DNA Res. 7 (1), 65-73 (2000)
20181126
2 (bases 1 to 5139)
Ohara,O., Nagase,T. and Kikuno,R.
Direct Submission
Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdna1nfo@kazusa.or.jp.
URL: http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)

```

```

FEATURES
source
CDS
gene
FEATURES
source
1..5139
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="fh11767"
/issue_type="brain"
/clone_lib="pBluescriptII SK plus"
1..4417
/gene="K1AA1312"
<1..4417
/gene="K1AA1312"
/note="start codon is not identified."
/codon_start=2
/product="K1AA1312 protein"
/protein_id="BA092550.1"
/db_xref="gi:7242979"
/translation="NSEHTAVISLCSGMLGTFRSHDGDYFTEPLQSMDEDEEONK
PHIYRARSAPQREPTSGHACDTSEHKNRHSKDRKTRARRKMGRIINAGDVALNSG
LATEFASVGNKTDNTPREKRVNRRKRELYPRVEVLVAVDNTMVMVYHGHNLOHYL
TMSIVASTYKDPSTGNINIVYLVNLIYHNEODGPTISENAOITLTKNFCOMOSKRS
PQGHNHDVAVLTLTKDDICRANHCDDTGLABELGTCDFRYSGSISESGSLSTATIH
ELGVFNMRHNDNNKCKEYKSPQHWAPLPIFYGVQVMSKCSKRYITFELIDVIG
ECLLNEBPSRPLVPLVPLGILYVNNKCELLIFGYSQVYMMQCKRMLNWNVNG
KGCQHTPMADGTECEBPKHCKYGVKEMADYDVLDSGMSYSPGTCSTRTGGGI
KTAIECRNRPKNGKGYKVGRRKMKFKCNTPEKQKRDREDDQCAHFDKHNING
LLPNVRYWYKSGIIMDKRCKLFCRVAGNTPAYLDRVLDGTPGDDINDICVOGIC
RQAGCDHYLNKARADKCGVGGSSCKYVAGTPEYVYVAVLVPYIPGATINDVR
OHSEGETDDVDYLAISSKGEFLNGLVYVYAKREIRINAYAYEYSGSEPTAYERN
STRLEQELLDVLSVGLYNDVAKYSINLITEDPQDFINNSHGWACSKPQDGER
KRKLVCTRESQDLVSDQCDRLPQPHITPECGTDCDLRWHVNSRSESKQCGLGR
LTDIYCAKYSRLDGTKEKYVDGFCSSHPKPSNRSCECMTGWRYSAMWTECSGD
GGTORRAICVMTNMDVLDSDSKCTHOEKVYTORCSEFPQMSGSDSVLTCGKCH
KHROYVCOFGEEDRLDMCDEFTKPTSMOTQOPEFCASMOAGPMSGCVTCGOGYOLR
AYKCTIGYMSVYDNDNCAATRRPTDODCELPSGHPRPADETRRSTYSAPRTQWR
GWTFCSATCGKGTMRKYVSCDENGSYADDSACTLPRYAKRECSYVTPCGOMKALD
WSSCSVTGQGGKATQVWCVNSDHYIDRSBEDDYIETEDQDSMSFCPQRTDSEGL
AQHPQNEDEYRPRRSASPSRTHYLGNQMTGPMWACSSSTGAGGQSRVYVQDIBNGTL
ANDCVERIKPDEQACESGPCPOMAVYGMWGECTKICGGGITRILVYCCORRNGEPRDL
SCETLIDKPRDEQCNTHACPHDAAMSTGPMSSCSYSCGRGKORRVYCMADGSHLES
DYCKHLAKPHGRKCGRGRPKWAKAGVSCSVSGRGOORHVGCCOIGTCKIKARETE
CNPYTRPSEPCQSPRRPLIYTRAKEMQDEECTKCTGSEBSRTRKYVVCVNDNKNVEVHGAR
CDVSKRPVDRSECSLQPEYVWITGEMSEVPSWEL"

```

LADPAASAYKTRDNRREKTRRTRKRLPSRPREVEVYLVADNRVYSHGENLQHYTL
 TLMSTVASYIKRPSSTGNLNIYIVNLVYHNEQDPSISEFNQOTPLKNEFCWQHSKNS
 PGGIHHDIYVLLTRDQICRAHDKCDPLTGLAEIGTICDPRSCSISDSDGKLNSTAFPIAH
 ECLHVEFMHDDNNKCKEYVSKPQHVMAPLNFTYKPMWMSCKSRKKTITELDLDTGYS
 DQILNEBERPPLVYOLEPGLLYNNKOCCELLFGBSOVCYPMQCRHLMONNVAHV
 KCRQHPHFWADGTEGEPGKHCYKFCVAKEMDVPVITGSMGSMSPRSTCSSTGGGI
 KRAIECRNPERKNGKYYVGRBKFKNSCTEPCIKQKRDREDECAHFDGHNFINIG
 LLPNRWYFKYSGIILMKDKCKLFCVAGNTAVYQLRDVYLDIDTFCGDTNDICVQGLC
 ROAGDQHVNLNSKARLADSSKGFELNNGVNTYMARERIGNAVYVYRIPAGATINRIN
 QHVSFGEEDDDNYLALSSKGFELNNGVNTYMARERIGNAVYVYRIPAGATINRIN
 SYDRIEOLLIQVLSYKLYNDVYRNFNIPJEDKPOOFLVYVPAKESRSTYPCGQKALD
 KRKLVCRESQDLVYSDORDBRLPQGHITRECGDQDLRHHVARSBSQCCGIGYR
 TLDICAKYKRDGKTEKYVDGFCSSNHRKPSREKCSSECVNTGKRIYAMTECSKCD
 GGTORRAICVYTRNDVLDSDSKTQHEKVTIDRSEFPCPOKSGDWSBCLVTCQKGR
 KHRQVCOQFEDRLNDRMCDEPTKPTSMQTCQAPCASWQAGPWCQCSVTCGQYQLR
 AVKCIIGIYMSVVDNDCNAALRPTDQDELEPSCHPAPAEETRSYSPAPRQWRF
 GWTPTCSATCGKGTMRKYSCDENGSVADEBACATLRRPVAKESSRSTYPCGQKALD
 CDPYRPRBESDCCQGRPLYTWRALAEWQECSTKTCGFEESRYRKKVYVCDNKNVEVHGAR
 CDSKRPVDRRESCLQPCPEYVWITGEMSEVPSWEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t
 ORIGIN
 AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..
 Initial Score = 5 Optimized Score = 11 Significance = -0.42
 Residue Identity = 18% Matches = 16 Mismatches = 50
 Gaps = 21 Conservative Substitutions = 50
 Translation Frame = 3
 TPSTRSSASAOQCMASHSGIMGIIILNHYSLMNMKMKRNKNTPTSTFGAARPRRSPQEGMHTVTPQNTKI
 10 20 30 40 50 60 70
 GTVTRRKRPEOENGERGLTWLVYXQHXTAAXQORHFLWYIRRTQEKRGKPRGQWVYFPIHGIKXKSMWQI
 80 90 100 110 120 130 140
 TEMPHMERTFNITIKYLCQDLKSLKTOVLEFIKLLILLLXKIFIMNRMGLPYLIMLRQHXKTFASGSIRR
 150 160 170 180 190 200 210
 TVQVESIMILLFEXODRISAEIETTVNIPYAMLNMEPFIPIEAVILVKKIVDYVQLLRSPMSWAMCITCLMMT
 220 230 240 250 260 270 280
 TTWVKKKELRVPKSMWLOHXKSTPTPTGCGSOVVENISLSEFKTLVMAVSLTQNLNDDPTCLISNCOASPTTXI
 290 300 310 320 330 340 350 360
 NNWVXFLDOVLRCAHIXCSADSGAITSMEYTKAAGLSTHPGPMGRSASLESTASMDVFFPKMMSPXOMDP
 370 380 390 400 410 420 430
 GEVGVLELPRAPREHVEGASKQPFESATDQOKWVENYVXVEKNLSPATRSVSHVSSSESEMSNVYLLIGSIL
 440 450 460 470 480 490 500
 TSTVCPMCAQGLSINVEFYKXKRGASCSGASAMQSTQPTISETEXMELLVARTOMISVSRFAFGKLDALIMFXI
 510 520 530 540 550 560 570
 OKFGEINVGFEVVAIILHAKOMOEHLQYIMVITLIMSEFQVLVPLIMCGSSTVSOGKQYMTTXYLXQAVKVNCS
 580 590 600 610 620 630 640
 X QLP-----GIL-----YVWNKOCCELLIFGPGSOVCPYMMQCRRLWC-----NNVNGVHKKGCRQ
 10 20 30 40
 XMETLLSQWPKGKFAFGMLMXSTVGRPLRXPXKELTQOIALSKNFCRFRQWESCTTPIVAVILISIPDLKTNLSS
 650 X 660 670 680 690 700 710 720
 HTPWADGTCEBEBGKCKYGFV X
 FT-----GTVMGSHKHNVAVNDPAKNGNENLFAPNLISLLFLIKDAITGCPSLDILLNRPVVOQTVYXGGMILPAGVN
 730 740 750 760 770 780

VVPSVAWVTAHNTSTVPRNLAGMGRRLRLMLMVFAAADPNQATVKNAGQNTVRVAGAILPRLGNVOKAVTVGR
 790 800 810 820 830 840 850 860
 GGGLEFVSIPEMMYWMTANHITKRIKLPFGAASVSLVHSGNLEFGOSAMSRYVEKGISTARSSVSLYKIKDMLEIC
 870 880 890 900 910 920 930
 VTLRPSQHLRLVSSRRNHPGRVRRVBDASVSLVDRDTSXEDQXNASLGLICOMXMTMVLWQOOLDLIDRPTVNY
 940 950 960 970 980 990 1000
 HHVILLPOLRKRKGEAHTHQEBSGDLGLPNAQPLVGVKVPKIDPSAAMRMAIMETRVPVLPICDDQORKNV
 1010 1020 1030 1040 1050 1060 1070
 LKHVVGNGRMTGALALXPLVGVKVGQPKXCVSTVTVTXSIGVYTRILISOKLITRTVPCNHALKRPGQVAXLS
 1080 1090 1100 1110 1120 1130 1140
 PPSKRRFTIYPGAAPAABMCSYETSSELAPGPHVVPVLDAPRSGVLLVYRMAKMDPTQTTVWREKMLMSKEYV
 1150 1160 1170 1180 1190 1200 1210 1220
 NPLVLSGLMATGESALSCVVEAXBQDMWVSVSGPTVNGFOIXAVKFLINLPIVSSVTHMLVHTTLLHGVLAIG
 1230 1240 1250 1260 1270 1280 1290
 ARLVLSVYEGINNEMETAMOKMEALIXKYITVSTWLSHMGTESABEEDAENKLAIGVSAICPVAEAYSRGMW
 1300 1310 1320 1330 1340 1350 1360
 AVRSHTKXPERPSATHTPPDNRNATYAKAHGVPSITLGGQRNKNAPRPAAKAPGTARNCVMMTTKTRCOMGHA
 1370 1380 1390 1400 1410 1420 1430
 VTXASGRWTVKAVVCPNPAMSGSGSONGQRYRPGNCHRHQLSHGLRVAGYVYEGXVQVOCGRLEATRPICHGEL
 1440 1450 1460 1470 1480 1490 1500
 CVFQOVLSEPKLEFLVCAAGEMXEMVSKSLQVYIIXFTLMSYMKKEIITGANSXDVAVRLKTVYKXKILX
 1510 1520 1530 1540 1550 1560 1570 1580
 COGYNKERWOLXQSTCCVKAESLISQILMPCRNVVXRYQYXFFKRGKSGLVCEISXLENGGXNPFEXNIV
 1590 1600 1610 1620 1630 1640 1650
 RQKQTPFNQMCAGLITNPFGRKXKXFSIISEVQYPPFRHISCCACQCSIRPPLAVIANDXXS
 1660 1670 1680 1690 1700 1710



KRAKTCRESDDIIVSDQRCDLRIPFGHITTEPCGTDICDCLRMHWASRECSAOCGLGAYR
 TLDIYCAKYSRLDGTKEVDKFCSSHPKSPNREKSGEBCNMGMRYSAMTEPCKSCD
 GGTORRRRAICVNTRNDVIDDSCCTHOEKVTIORSSEFPCQKSGSDMSFCIALYTCQKGH
 KHRQVWCGEGBRLNDRCMDPERKPTSMOTCCOPBCASQADPQWCGSYTCGQIQUR
 AVAKLIGITMIVYDDNDCHAATRPDITQDELBPSCHPRAPAERSTYSAKTRKQ
 GSWTPCSATCGRATRMRYVSCBDENGVADESACATLPPVAKKESVTPCGOMKALD
 MSSCVTCGGRATQVQVYVNDVHIDRESDODVITPDDDCSSPCPOTPTGSL
 ACPHONYERYRPPASASRTHVYGIMQRTGFMGACSTTCAGGSOBRVYVCGDDENGYT
 ANDCVERTIKIDDEBACESGPPQMAVGNMGECTKILGCGGIRPLRYVCCORNGERPDI
 SCPIILKPPREDCCNTNACBPDAAMVSTGWSYSCSASCGGGHQRWYKCAKRSKHLES
 DYCKHLAKPHGHRKCRGRCPRKAKGAMSQCSVCGPRYQVRHVGQIGETHIARETE
 CNPYTRPESREDCQGPRLTYTWRAEWQECCRTKTCGEGSYRKRKYVVCVDNKKNEVHGAR
 CDVSRVYVRESVSLQPCREYVWTGEMSEVPWEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t
 ORIGIN
 AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..
 Initial Score = 129 Optimized Score = 129 Significance = 2.05
 Gaps = 0 Matches = 129 Mismatches = 0
 Translation Frame = 2 Conservative Substitutions = 0

EMGSPDAANVKRDLHPROYKLLPETSSEYEVSPRIYVNLGEPFPPNPNVHFKRTSRINSATSATDPWPAPFASSS
 SSSTSSQAHYRLSFAFGQQLFNLNTANAGFIAPLFTFTVLLGTPGVNQTRFYSEEAELKHCFYKGYVNTNSEH
 150 160 170 180 190 200 210
 TAVTSLGSMGLGTRSHDGDYFIPLQSMDEDEEONKRRPHLIYRRSARFORRESTRHNACDSEHNNRHSK
 TAVTSLGSMGLGTRSHDGDYFIPLQSMDEDEEONKRRPHLIYRRSARFORRESTRHNACDSEHNNRHSK
 TAVTSLGSMGLGTRSHDGDYFIPLQSMDEDEEONKRRPHLIYRRSARFORRESTRHNACDSEHNNRHSK
 10 20 30 40 50 60 70
 DKKATRRARWGERINLAGDVAVALNSGLATEFSAVGNKTTDTRREKTRHRRTRK
 DKKATRRARWGERINLAGDVAVALNSGLATEFSAVGNKTTDTRREKTRHRRTRK
 DKKATRRARWGERINLAGDVAVALNSGLATEFSAVGNKTTDTRREKTRHRRTRK
 80 90 100 110 120 130 140
 SYGENLQHYLLTLMSTIVASLYKDPSPISGNLNIIVLNILVIHNEDODSPSIFNAQTTLKNKFCOMOHSKNSPG
 150 160 170 180 190 200 210

GIHNDTAVLLTRDIDICARHDKCDTLLGELACTIDDPYRSCSISESDSLSTAFGLAHNELGVFPMRPHDNNK
 230 240 250 260 270 280 290
 KEEGVKSPQHVMAWPLTFEYTPNPMWSKCSRRKYITTEFLDTGVECLNLEPESRPYRPLFVOLPLGILVYNNKQCE
 300 310 320 330 340 350 360
 LIFPGSGVQVYVVMQOCRLMNNVNGVHNKGCRRDHTPMDAGTDECEPKHCKYGCYKPEKMDVAVDTGSGWGSW
 370 380 390 400 410 420 430
 SPFGTCSRTGGGIKTAIRECNREPEKNGKCYCVGRMRKFSQNTPEPCLKOKRDFRDEOCANHFQKGFHNING
 440 450 460 470 480 490 500
 LLPVWRVWPKYSGLIMKDRCKLFGRVANGNTAYYQRLRVDYIDGPPCGDVIYDIOVGLCRAGCDHVLNSKAR
 510 520 530 540 550 560 570 580
 RDKCGVGGDSSCKTVAAGTEFNIVYGVNTPVIRIPAGATINIDVROHSFSGEETDDNYLALSSSKGEFLLNGN
 590 600 610 620 630 640 650
 FVYTMARREIRIGNAVYVESSEFAVERINSTRIEDELLOLVISVSKLKNPDRYRSFNPIEDKKEQOQRFYVN
 660 670 680 690 700 710 720
 SHGPWQACSKPQCEGERRKLVCTRESDDIIVSDQRCDLRIPFGHITTEPCGTDICDCLRMHWASRECSAOCGLG
 730 740 750 760 770 780 790
 YRTEIYCAKYSRLDGTKEVDKFCSSHPKSPNREKSGEBCNMGMRYSAMTEPCKSCDGGTQORRRRAICVN
 800 810 820 830 840 850 860

TRNDVIDDSCCTHOEKVTIORSSEFPCQKSGSDMSFCIALYTCQKGHKHKHROYWCGEGBRLNDRCMDPERKPT
 870 880 890 900 910 920 930 940
 SMOTCCOPBCASQADPQWCGSYTCGQIQUR
 950 960 970 980 990 1000 1010
 APRTRSTYSAPRTRQWRFPSMTPCSATGCTGRMYVSCRDENGSVADESACATLPPVAKKESVTPCGOM
 1020 1030 1040 1050 1060 1070 1080
 KALDMSGCVTGGGGRATRQVWVSDHVHIDRESDODVITPDDDCSSPCPOTPTGSL
 1090 1100 1110 1120

2. US-09-972-467-2 (19-287)
 ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
 TOIG of: ab037733 check: 4803 from: 1 to: 5139
 TOIG of: ab037733 check: 4803 from: 1 to: 5139
 LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
 DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
 ACCESSION AB037733
 VERSION AB037733.1 GI:7242978
 KEYWORDS Homo sapiens brain cDNA to mRNA, clone_lib:pbLuescriptII SK plus
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 clone:fh11767.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo.
 REFERENCE 1 (sites)
 Nagase,T., Kikuno,R., Ishikawa,K.I., Hirotsawa,M. and Ohara,O.
 Prediction of the coding sequences of unidentified human genes.
 XVI. The complete sequences of 150 new cDNA clones from brain which
 code for large proteins in vitro
 JOURNAL DNA Res. 7 (1), 65-73 (2000)
 MEDLINE 20181126
 REFERENCE 2 (bases 1 to 5139)
 Ohara,O., Nagase,T. and Kikuno,R.
 Direct Submission
 Submitted (31-JAN-2000) Otsamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
 292-0812, Japan (E-mail:cdna1nfo@kazusa.or.jp
 URL:http://www.kazusa.or.jp/nudge/, Tel:+81-438-52-3913,
 Fax:+81-438-52-3914)

FEATURES
 source
 gene
 CDS
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="fh11767"
 /tissue_type="brain"
 /clone_lib="pbLuescriptII SK plus"
 1..4417
 /gene="KIAA1312"
 <1..4417
 /gene="KIAA1312"
 /note="Start codon is not identified."
 /codon_start=2
 /product="KIAA1312 protein"
 /protein_id="BA942550.1"
 /db_xref="GI:7242979"
 /translation="NSEHTAVVISLCSGMLTGFSSHDGDIYIEPLQSMDEDEEONK
 PHLIYRRSARFORRESTRHNACDSEHNNRHSKPKKKTTRAKWGERINLAGDVAVALNSG
 LATEFAVAGKTDNTRREKTRHRRTRKFLSYPRFVYLVADNRNYSYFNENLQHYIL
 TMSIVASIKYKDPISGNLNIIVLNILVIHNEDODSPSIFNAQTTLKNKFCOMOHSKNS
 PGIHTDPAVLLTRDIDICARHDKCDTLLGELACTIDDPYRSCSISESDSLSTAFGLA
 ELGVFPMRPHDNNKKEEYVKSQHVMAWPLTFEYTPNPMWSKCSRRKYITTEFLDTG
 ECLINPESRPYRPLFVOLPLGILVYNNKQCELEFGPSQVCPRYMOCRRLMKNNVNVG
 KGRTOHTPMDAGTDECEPKHCKYGCYKPEKMDVAVDTGSGWGSWSPGCRRLMKNVNVG
 KTAIRECNREPEKNGKCYCVGRMRKFSQNTPEPCLKOKRDFRDEOCANHFQKGFHNING
 LLPVWRVWPKYSGLIMKDRCKLFGRVANGNTAYYQRLRVDYIDGPPCGDVIYDIOVGLC
 ROAGCDHVLNSKARDKCGCGGDNSSCKTVAAGTEFNIVYGVNTPVIRIPAGATINIDV
 OHSFSGEETDDNYLALSSSKGEFLLNGN
 YRTEIYCAKYSRLDGTKEVDKFCSSHPKSPNREKSGEBCNMGMRYSAMTEPCKSCDGGTQORRRRAICVN
 STRIEDELLOLVISVSKLKNPDRYRSFNPIEDKKEQOQRFYVN

KRKLVGTRSDOLTVSDQRDLRPFQGHITTEPQCTDCLDRMHWASSEGAQCGLGR
 LLDIYCAKTSRLDKGTEKVDGDFCSHPPSPNSKCSGRCNMGGMWYSAWTECSKSD
 GGTQRRRAICVNRNDVLDSDSCCTHQEKVLTORCSFPCQPMDSGDMSECLVCGGCH
 KHRQWVQGFGEIENDRMCDPEPTPTSMQTCOPPECAQWAGMAGCCVCGGCGYDLR
 AVKCLIGTVSVVDDNDCAATRPPTDTCOCELSCHPRAAPGRTSRSTYSAPPTQRF
 GSWTPECSATCGGRTKRMVYVYSCDRENGSVADSEACALPRLVAKRECSVPPGQMKLD
 MESSAVTCCGGGRTKRMVYVYSDHYIDRESCODDITPEDDDCSMSPQRPDSDGL
 AQHPNEDYRRRSASPRSTHVLGQNRTRTGPAGCSSTGAVGQRRVVVQCODEGT
 ANDCYERIKPDEQRACESSPCQWAVYGNWEGEKLLCGGGLRTLLVVCQSRNGRFPDL
 SCEIILDKPRDREQNTNACPHDAAMSTGESSCSVSCGGRHAKORNVYCNAMKDDSHLES
 DYCCKHLAKPHGHRCKGRCPKWKAGAMSGSYSCGCGVQOQRHVGQIGIENKIKARRE
 CNBYTREREEDCGPRCPLTYTWRAEMOECTTCOEGESRYRKYVVCDDMKNEVHVAR
 CQVSKRPVRESCLQPCHEVMTGEMSEVSMEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t
 ORIGIN
 AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803
 Initial Score = 11 Optimized Score = 47 Significance = -0.34
 Residue Identity = 208 Matches = 63 Mismatches = 189
 Gaps = 58 Conservative Substitutions = 0
 Translation Frame= 4

MTTGLPLKRVXISIVHRIKYEKMLEILRLKIIIFFLKSCLLNRHNTSGSKYKGFVCPLOCQPKTKMFSPH
 10 20 30 40 50 60 70
 FOYRKSIIIVQIFSLSKKIRSGNARPPRCMAATAGVKEIILPDSHSMCKYFVVAISPCIPGTTITSYIXLSSG
 80 90 100 110 120 130 140
 TSOHPMSLOPLXKFLVLANRTSRXITTECRREPTXPSLITSLPCLICIGILAMESXGLVYRKHHTVHSGKSGQXPQ
 150 160 170 180 190 200 210
 VSHIPLILPRLPRLSGHGKADDDGSSQDSTSPHSVVIQYISOGCKLQLSRNGRLLITSORACTSIFLSS
 220 230 240 250 260 270 280
 THHTTLRLVLEPPOVLVHSHSSALQVXGHRGHPWQSDSGLVYGLHVSALAILCVPIQXPTCLCTPRRPOD
 290 300 310 320 330 340 350 360
 TEHXLQAPAFHLLGHLPRPHFLCPCGLARCLQXSLSKWLPSPAMQXTRPCLCRPQTEPHLELQGVLANASGC
 370 380 390 400 410 420 430
 QACVLCRSGRGGLSRISQKSGNRSPLDRXQYQVTSVLMPPRHSLVHSPDLFXHXHGQGPDSQALCSSGFTILS
 440 450 460 470 480 490 500
 TOSFAVYVFPSSXHTTTRRWDPRQVLELHARQGVYLVHWPPTSTWYRILGALURGRSSPWWKGGXAKPDLISGLXG
 510 520 530 540 550 560 570
 HGDMEQSWSVSGIXSMWSHLRSTITWSLXLNHTGCRVALPCQVTEOELQSKAFHCPOQVTEHSSPATGLGRV
 580 590 600 610 620 630 640
 AVALSSATPEPSSRQLTYLLIRVPRQVAEHEGVODPNNRHVYLIGALVYLLVSGAAGGKHDGNSQSWVSYGLV
 650 660 670 680 690 700 710 720
 AALQSSSTTDIXVPMHMTALSWYPCQVYTLHRCQSPACODAHSGCXOVCIDVGLVSGSHILLSFNRRSPXK
 730 740 750 760 770 780 790
 HQFRCLCPRPQVYTKHSDQSPDFHCQGGNSLHLXMTVEFSXVHLLSSSTSRVLTJOTALLLWVPPSOLFHS
 800 810 820 830 840 850 860
 VQAEKRPVYVHSEPHSRLLGLGWLIDQKPSSTFVPSLLYLAQXMSVNRKPKRMWALHSLDLATCTHRS
 870 880 890 900 910 920 930
 X EMGSPDAANAIVRKRRLRHRQYKLETLT--SEYELVSP--ITVNAALGEP--PPTN-----VHKRRPRR
 20 30 40 50
 OSVPOGQVGMKPRGWSRHLXSETVYSXSDSLVQYTSFRFRSPWQGLLHACHGXYLQXCMXGSLSTIGLKEKXT
 940 950 960 970 980 990 1000
 S--INSATDDPMPAFASSSSSTSSQAHYRILSABEQDFLNFITANAGTAPLPTVTLGCTGCVNDTKR-----
 60 70 80 90 100 110 120
 IGLVNFPTD---KTXSSKCSMRS-VELLISTAVSDPLYSTTAPRMKISLAIYV-----TKRPFPSKN
 1010 1020 1030 1040 1050 1060

230
 140
 2817
 269

-----YSEBEALTKKCFKYGWVNTMSEHTAVILSCGMV---GTRSRGDYFIIPQLQMDQOE--
 SPLLDKRAKXLLSSVSPKCLCCFRSILVAPAGITFTVYXPKCTVLANPAAVTLHEBELSPROPHLSIRAFERK
 1070 1080 1090 1100 1110 1120 1130
 180 190 200 210 220 230 240
 ---DEEONKPHIYYRSABQREPRSTGRHACDTSEHNKRHSKDKKTKRARKWGERINL-AGDYAALNSGLA
 TXSHAPCRQRPWTQISLPSWQGVPS---TILSRXXXXAVPRTLQNNLHR-SFRIRPFLYIGTORLTSRRL
 1140 1150 1160 1170 1180 1190 1200
 250 260 X
 TEARSAVGNKTKDNTREKTRRTRKR

MLKCFPSKXAHSSRRKSRFCRLRHGCSVLQDLNFIILRPNQYPPFPFGSGLLHSRMVAVLMPPHVLEQVPRKQLD
 1210 1220 1230 X 1240 1250 1260 1270
 PDDPSVMTGTSISLGTQNPYLQCFPGSHVPSAQGVYXRPDLCTPRLILLHQSRLHCITYGHTXEPGPKNSH
 1280 1290 1300 1310 1320 1330 1340 1350
 CLFTLXRMPSGWTGKXGLDSDGSLSKSHSPXPKNSVLYYRFLHFDHLOGLVXKFSVCGAMTCWGLLTPSSLH
 1360 1370 1380 1390 1400 1410 1420
 LLSXGMILNTPSSWALVYKAVLANPLSLIEQLXGSGQWVPSARPKYSHLSXALQISCLVYKRTVAVSKWIPPG
 1430 1440 1450 1460 1470 1480 1490
 LFPECCHWQKFNVYVXALKDMEGSPCSLXITIKFTITLILKFPILGISLXIPATIDIKYKIXCFRSPWYETI
 1500 1510 1520 1530 1540 1550 1560
 LLSATYKTSRNRGDKRRKRFVILWVLFSLVSVLLPXAENASVAPRLFNAAATSPRLLISPHFALVFLSLL
 1570 1580 1590 1600 1610 1620 1630
 CLFLCSEVNSHCLPDEGSLWGLRDLXMMWGLDFCSSSSSIDCSGSIKXSPSXDRNVPSPIDORLMTAVCS
 1640 1650 1660 1670 1680 1690 1700 1710

3. US-09-972-467-2 (19-287)
 ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
 TOIG of: ab037733 check: 4803 from: 1 to: 5139
 LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
 DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
 ACCESSION AB037733
 VERSION AB037733.1 GI:7242978
 KEYWORDS
 SOURCE Homo sapiens brain cDNA to mRNA, clone_lib:pluescriptII SK plus
 ORGANISM Homo sapiens
 clone:fh11767.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Nagase,T., Kikuno,R., Ishikawa,K.I., Hirotsawa,M. and Ohara,O.
 TITLE Prediction of the coding sequences of unidentified human genes.
 XVI. The complete sequences of 150 new cDNA clones from brain which
 code for large proteins in vitro
 JOURNAL DNA Res. 7 (1), 65-73 (2000)
 MEDLINE 20181126
 REFERENCE 2 (bases 1 to 5139)
 AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
 TITLE Direct Submission
 JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology: 1532-3 Yana, Kisarazu, Chiba
 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
 URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
 Fax:+81-438-52-3914)
 FEATURES
 source 1..5139

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="fhl1767"
/clone_lib="brain"
/issue_type="brain"
/clone_lib="pbluescriptII SK plus"
1..4417
/gene="K1A1312"
<1..4417
/gene="K1A1312"
/feature_start="not identified."
/codon_start=2
/product="K1A1312 protein"
/protein_id="BAA2550.1"
/db_xref="GI:7242979"
/translation="NSEHTAVILSGMLGTFERSHDGDFTEPLOSMDREDEEONK
PHIIRRSAPORREPTGRHACDTESEKRRHSKSKKTRAKKMERINLNGALNALSG
LAFEPAYAGNKTNDPFRKTRHRKREPLSPYVAVLADNRAMVSYGNDLONHTL
TMSIYASIKDPSIGLNIYIVNLIYHNEEDGSIENNOTLTKRNCOWSHKNS
PGGIHNDIYVLRQDICRAHDKDVTGLAELDTIDPVRKSCISIDSGISTAFITAH
ELGVFMPPHDDNNKCKEKGVSPOHVMARPLTLPNPMWMSKSKRYITEFLDVTYG
ECLINERESPYPYVQLPGLIYVNVKOCLELIGPSSQVRYMAGCRLMKNVGVH
KGCRTQHTPMADGTECEPCKKXKCYCPKEMDVPYTDGSMGSMSPFGCSRTCCGGI
KTATRECNREBPCKGKCYVGRMKKEKSCNTEBCLMOKRDFRDEOCADHDKHFNIN
LIPVRRVPRKSGIKMKDRCKLECAVAGNATAYQLDRVITDTPCCQDIPNDICVQCL
ROAGCDHVLMSKARDKCGVGGDNSSCKTVAETFNIVHIGTVVIRIPAGATINDIVR
OHSFSGEITDDNYLALSSKGEFLNNGNFWTVAKRRIKGNVAVVEYSGSEIYVERIN
STRIEDELLOVLYSGLYNPVRYSFNIPIDKPOQFPMNSHGWQACSKFCQGER
KRLVJCRESDQLVSDQCDRLPQPHITEPCTDILMHWASRSCSAOCCGLXR
TLDYCAKYSRDLQKTEKVDGFCSSHPKSNKESGECNMGGMRYSAWTECKSGCD
CQGNRRATIGVNRVYLDKSCDHOEKVITQCSERPCQKSGSMSCILYTCGCGH
KHQVWCOFEGEDNDKCDPRTKPTSMQTCQDPECSWAGWGCVATCGOQIQR
AVKICITVMSVVDNDNANATPTDQDELCECHPAPARTRSTYVAPGQWIKR
GSMPCSATGKGTMRVYVSRDENGSVADBSACATLPRVAKECSVTPCGQMKALD
WSGCVTCCQGRATQRYMVCVYNDHVIDRSECCQDYIPEIDQCSMSPCQRTPDGGL
AONFVERIKRDEORACGSPGQMAVGMGCEKIKLGGTGRFRLVYVCSNSNGRPPDL
ANDCCVERIKRDEORACGSPGQMAVGMGCEKIKLGGTGRFRLVYVCSNSNGRPPDL
SCETLDPKPRRBCNTHACPHDANWSTGPPSSGSCGGRHKORANYCAKDSHLES
DYCNHLAKPHGRKCRGRCPKWAKAGSQCYSVSCRGVQDRIHVGQIGTKIARTEP
CNPYTRPESRDCQGRCPLYTKRAEMOQCTCTCGEGSHYRKYVCVDDNKNKMEVHAR
CDVSKRVRDESSSLDPCFYEMWITGSEVSPWEL"

```

```

BASE COUNT      1471 a      1112 c      1362 g      1194 t
ORIGIN
AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803

```

```

Initial score      = 8      Optimized score = 43      Significance = -0.41
Residue identity  = 19%     Matches         = 60      Mismatches    = 175
Gaps              = 72     Conservative    = 0
Translation Frame= 6

```

```

          X      10      20      30      40      50
EMGSPDAAAARVRRKRLRHRQVQLLETLSEYIYSPTRVYVALGEPPTIVH
DYWYFTTIKKGIDKRCQAQEDIKKNGTQXTSDIENYFL--PKL-FYKQNDHFMKGV-CLRLTMYXNK
          10      20      30      40      50
FKRTRRS-INSATDP-----WPAFASSSSSSTSSQAHYRLSA-FGQ-----QFLFNLTAANG
FQRPSSKSEIHSITSPDFPKLKNQWXXKXKTFELHGHMS-----ERDSAFPQTHVLVCHSCHLSLILYPMWY
          70      80      90      100     110     120     130
FLAPLFTVTLTGTGPGVNTKIFYSEEA-----ELKHCFYKGYVN--TNSEHTAVILSGMLGTFERSHDG-D
NIFLVLVYFRRHJT-----ASYEFAPRVIISSICQXDIKVNHRMXYXRDLLEFIVSYSPALHTRNSISLGL
          140     150     160     170     180     190     200     210     220     230
YFIEFLQSMDEQDEDEEONKPHIYIRRSAPQREPTGRHACDTESEKRRHSKSKKTRAKKMERINLNGALDV
RRCCKTHSSOMQIGRVASSLRH-----SCTNHPSTHRSATLRRPWLKSRWQVQFGRVILXPRSCDPPDILA---
          200     210     220     230     240     250

```

```

240      250      260      X
AALNSGLATEAFSAVGNKTDNT-----REKTRHRTK
-----GIQTAFVNHRYRLAHVHTACRMLHLYVVYTHNHLAVRQAFAGLGFAPRLICRPSVEGTPMALAVAFR
260      270      280      290      300      310      320
LRGWWVVLGLSGYFVCSDDLAMRPLLYASANGHRAALRPSASEPFGASSSALSVPMMLSOVLTVTITFEMAS
330      340      350      360      370      380      390
IFGHANVLSLPMSTTRDRTRARASTPQSVWYVSMCYLLTLTGRIKNTAOKMKRFTYVGGPLTDHQSCSYA
400      410      420      430      440      450      460      470
STTOLSALEVAISPRLTRAGFTGDLIRFYSLHTVCGVSIFFILTYNNTPLGASSTGTCSPGASPELVS
480      490      500      510      520      530      540
TEHMAAGANAPRTYLLILEGVLSTAKTVMGRLRAWHXGTVLVSFMDIILLVTLPIIDHVVTVVDTHNLRPCTR
550      560      570      580      590      600      610
LPTGHRARAEVQGLRPLPTGCHRTFFLCHWSROGSTGTLVSHRAILISAADVSHRPGTFPTSGXAWGPRPKSPL
620      630      640      650      660      670      680
GSMCTVQASPRFRGSMGRMTXKXETVLGISMSSCCITVIVYHXHISPNDAFHCSXLSLSTSDTALSPTGR
690      700      710      720      730      740      750
LPCSTFRLLTSLHRQWLGRLVTHSIIIXSIFPKLTPDLAVLMPSTGDAQLXAPVSRPPLMTRBLAPLNGNFI
760      770      780      790      800      810      820      830
LMCAFAVIOYIISGIDTNSPSPRLPVTVAETXETSPGRVAPVATVPRXAPFTVAEPGMAAAKTIINLSLPI
840      850      860      870      880      890      900
QPAIFGVVDVQCAVYQATLGTETPTPAQNMPEQVYVCTGTFGSVNSRLQDRIASLIRNSKILIRPPKANKESEFP
910      920      930      940      950      960      970
PLAGPTACLIPWMTVPVKKLRFIFENNNIERIAYIGVQVLSHRQNLKQKFLNAICGVNSFYGSLGPTVLYHS
980      990      1000     1010     1020     1030     1040
IPMANPFEHGDKNKSIXOEFTEPTAAXSXVVVIVCFPRXETVLRPHINIGSTSNNSDHSIVTMYGCIKSCCHF
1050     1060     1070     1080     1090     1100     1110
AXRTIATNTNPTFSPGYYXNMTASSLPAKALDITDIIQVLAIRFSSSIYHSVSKLIYGVCPCHSABQDAPLVHQ
1120     1130     1140     1150     1160     1170     1180     1190
NSTVFRDPAHIGKQTVDKMLPVKSTLFTISEVSLLETWLRVAGLKEFHSTVYFSTIEMFWSVALNSGCF
1200     1210     1220     1230     1240     1250     1260
DAPSTGSGAGSKGTPSPGSIHGDHIFPFGNFK
1270     1280     1290

```

```

4. US-09-972-467-2 (19-287)
ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
TOIG of: ab037733 check: 4803 from: 1 to: 5139
LOCUS      AB037733          5139 bp      linear      PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for K1A1312 protein, partial cds.
ACCESSION AB037733
VERSION   AB037733.1 GI:7242978
KEYWORDS
SOURCE    Homo sapiens brain cDNA to mRNA, clone_lib:pbluescriptII SK plus
          Clone:Fhl1767.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites) Nagase,T., Kiyuno,R., Ishikawa,K.I., Hirosewa,M. and Ohara,O.
          Prediction of the coding sequences of unidentified human genes.
          XVI. The complete sequences of 150 new cDNA clones from brain which
          code for large proteins in vitro
JOURNAL   DNA Res. 7 (1), 65-73 (2000)
MEDLINE   20181126
REFERENCE 2 (bases 1 to 5139)

```


REFERENCE Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
 1 (sites)
 AUTHORS Nagase,T., Kikuno,R., Ishikawa,K.I., Hirosewa,M. and Ohara,O.
 TITLE Prediction of the coding sequences of unidentified human genes.
 XVI. The complete sequences of 150 new cDNA clones from brain which
 code for large proteins in vitro
 JOURNAL DNA Res. 7 (1), 65-73 (2000)
 MEDLINE 20181126
 REFERENCE 2 (bases 1 to 5139)
 AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
 TITLE Direct Submission
 JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba
 292-0812, Japan (E-mail: dnainfo@kazusa.or.jp,
 URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
 Fax:+81-438-52-3914)

FEATURES

Source Location/Qualifiers
 1..5139
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="FH11767"
 /tissue_type="brain"
 /clone_lib="pBluescriptII SK plus"
 /gene="K1A1312"
 <1..4417
 /gene="K1A1312"
 /note="Start codon is not identified."
 /codon_start=2
 /product="K1A1312 protein"
 /protein_id="BA92550.1"
 /db_xref="gi:7242979"
 /translation="NSENHVAVILSICGMIGTFERSHDDGYFIEPLQSMDEDEDEEONK
 PHIIYRSAPORPSTGTGHAACDTESEKRNRSKDKKTRARKMGERINLAGDVAALNSG
 LAEAFSAVGNKNDTRREKRTHERTRKPLSYPRPVLYLVANNRMVSHYGMLOHYLIL
 TLKSIYASIKKDSIGNLNIYVNLIVLHNODGDSIENNOTLTKKCGQMOHSHNS
 PGSIHNDVAVLTRDICTRAHDKCDTLGLAELDTIDPFRSCSISEDSLSLSTAFPTAR
 ELGHVEMPRHDDNNKKEEIVSPOHVMAPTLLEFYTNPMWMSCKRYITFELDLYG
 ECLINPESEPRPLPVQLPGLVYNKQCELLIFGSPQVCPYPMQCRRLMKNVNVGH
 KCGRTQHPADTECEPQGHCKYVFCVPEMVLVTKDQSMGSMSPFGSCSTGGGI
 KTAIRECNRPEKNGKCYVGRMRKSCNTEELCKOKHPDFDEOCAGHDFKHPINIG
 LIPNVRVPEYSGIILMKDKLFCRAVGNAAVYQLDSDRVIDTGPCODINDICVQGLC
 ROAGCDHVLNKRKRDKCGCGDSSCKTYVYQLEPMTYKYGATVYRTPAGAPNIDVR
 QHSFSGETDDNTLALSSKGEFLNENVTYMAKREITIGNAVVEYSGSEPAVERIN
 STRIDIEQLLQVLSVSKLVPDVRYSFNPIEDKQDFYMWSHGWAQCSRQCGER
 KRRLVCTRESDDQLTVSDORCLDLPQGHITPEQDGLRMHVAISSECSACCGGGR
 TLDIYCAKYSRLDGTKEKVDGFCSSHPKPSNREKSGECNIGMRYSWMTSCSXD
 GGTORRRAKCVNTRVNDVLDSSKCTHQEKYTIORCSFEPQPKSGDMSRCLVYCGKGH
 KHRQVWQREDELDNDRMCDPRTKPTSMOTQDQPERASQAGRWGQCSYTCGGQGTQLR
 AVKCIIGTIVSVVDNDCNATRTPTDIOCELDLPSCHPRAADPTKRSYTSAPRTQRE
 GSWTPCSATGKTRMYVYSCRDENGVADESACATLPPRVAKREKSVTPCGQMKALD
 WSSCSVTCGGRAITQVMVYNDHVIDRSECDQYIIPETDDKSSPCPQRTPSGL
 AOHFPOVEDYRPRASPSRTHVLGNGMRTGPMGACSSITCAGGSGORVVVVQDDENGYT
 ANDCVERIKRDEORACESGRCPOWAGNMGECIKLGGGIRIRIIVYVCOYCNKREBFDL
 SCEILLRPPRDECCNTHACPHDAKAMSTGPRSSVSCGCGHKRQWYYSQAKKGSHELS
 DYCKLHAKPRHRCRGRGWKAGAMSGQSSVSCYGGQQRHVVGQIQITNHRITARTE
 CNPRTRESEDCQPRCLYTWRAEWEQECTTCGEGSRYYRKYVVCVDDNKRKNVEHGAR
 CDVSKRPRDESSQLPCEVYVITGEMSEVPSWEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t
 ORIGIN

AB037733 length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ...
 Initial score = 7 Optimized score = 21 Significance = -0.43
 Residue Identity = 20% Matches = 32 Mismatches = 82
 Gaps = 41 Conservative Substitutions = 0
 Translation Frame= 3

EWGSPDAAAAVAKRDLRHPVAKLLETLESEYEVPIRANALGPEPPIVNHKRTRRINSATDPMWPAFASSS
 * 80 90 100 110 120 130 140
 SSSSTSSQAHYLSAFAAGQGFLENLANAGFIAPLFTVTLGLTPGVNOKTRFYSEEAELKHCFYKGYVNTNSEH

150 160 170 180 190 200
 TAVIS-----LCSG-MLGTFFRSHDDGYFIEPL-...OSKDEDEDEEONKPHIYRSAPORPSTGTGHAACD
 PPSASNOECWANSGLMMG-----IILLNHYSLWMMKMKRNKTNPTSFII--GAAPPRESPQDRCMHV
 10 20 30 40 50 60
 210 220 230 240 250 260
 TSEHKNRSHKDKKTRARKMGRINLAGDVAALNSGLATEA---FSYGNKTDTRKR-----T
 T--PQNTKIGTAVK-TR-RKPEDENGEKGLTWLVYQHXTAAAXOQNHFLMVTIRRTTQERKRPTEGQNVFPI
 70 80 90 100 110 120 130
 X
 HRRTKR
 HGLKXSMWVQTTWEPRTMERTKENTTEFXLXCQLPRLSIRKTVLEIXLILLXTXLFINRMGLPYLLMKRH
 140 150 160 170 180 190 200

210 220 230 240 250 260
 XKTFASGSTRRTVQVSEIMILLFSXQDRISAEITTNVIVPXMWLMNEPVIPIEAVLVKIVDYXQDLRSPMS
 280 290 300 310 320 330 340
 WAKCLTCLMNTTNNKKELELRVPSMSWLOHXSTPTPGCGQSVVENISLSEXYTLVMASVCLTNNLPDPTLCL
 350 360 370 380 390 400 410 420
 SNGQASFTTXINNVNKFILDVLRCAHIXGSADSGSAGATSMETKAAAGLSTHFGPMSASLESFTASMDPVP
 430 440 450 460 470 480 490
 KKMSPXQMDPGEVGLPEAPAEHVEGASKOPFESATDQNMOMVENYVYDXEKNLSPATRSHVSRSETSEM
 NSVLTLTGSLITSTVCEFPKAGSLNLTVEFXHRTGASCSAEMOQTPTISFTEEXXMEILLVARTOMISVSRAP
 500 510 520 530 540 550 560

570 580 590 600 610 620 630
 AGGLDLMEXYTKRGEINIVGFVYVAILLNAKQWQENHIDQIMVITLISEPQVLLPLLMGSGTYSGKKQMTTIT
 640 650 660 670 680 690 700
 XLYQAVKNSXJMETLISQMPKGFALGMIMXSTVGRPLRXKELTQOIALSKNFCRFRWESCTTPMAYAIL
 710 720 730 740 750 760 770 780
 SITQKLNLSSEFTGTVMGKRAVNPARKNGNENLFAFGNLLSLEFLIKDAIGCPSLDITLNPVYQVYTXGS
 MPRAGVNVVPSVAMVTAHNTSTVFNPIAGMWRRLRMLMFAAALIPNOATVKNKQGNVTRVAGAILDPLGNVOK
 790 800 810 820 830 840 850
 AVTVGSRGEGELFVSTEDMAYMTANAHIRKRLPRFGAVSSLVHSGNLEFGSAMSRYVEKGISTARSGVSLVK
 860 870 880 890 900 910 920
 IDXMIKCVTLRPSOHLICRLVSSRNHVRGRRVPGDSAVSLVDSDTSXEOXNASLGLICOMXMTMVMQOJDLQ
 930 940 950 960 970 980 990

1000 1010 1020 1030 1040 1050 1060
 IPTVTYVNHVILPDLPRKKGAEHATVQNEPSCDGLGSRPNAQPLVGVKPKVDTSAAEMKRALMVTTRPVLPCLD
 1070 1080 1090 1100 1110 1120 1130
 OMQRKRWLXHPVGNRGRPTGALALXLPVGVGVPGRKXCVSTVTVTXSIGVSVYRIISOQLTTRTPVPCNALK

6. US-09-972-467-2 (19-287)
 ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
 TOIG of: ab037733 check: 4803 from: 1 to: 5139
 LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
 DEFINITION Homo sapiens mRNA for K1A1312 protein, partial cds.
 ACCESSION AB037733
 VERSION AB037733.1 GI:7242978
 KEYWORDS Homo sapiens brain cDNA to mRNA, clone_lib:pBluescriptII SK plus
 SOURCE

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE Nagese,T., Kikuno,R., Ishikawa,K.I., Hirosewa,M. and Ohara,O. (1998) Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro
 JOURNAL DNA Res. 7 (1), 65-73 (2000)
 MEDLINE 20181126
 REFERENCE Ohara,O., Nagese,T. and Kikuno,R. (1998) Direct Submission
 TITLE Laboratory of DNA Technology, Osamu Ohara, Kazusa DNA Research Institute, 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp)
 JOURNAL URL: http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913, Fax:+81-438-52-3914
 FEATURES Location/Qualifiers
 source 1..5139
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="fhl1767"
 /tissue="brain"
 /clone_lib="pbluescriptII SK plus"
 /gene="K1A1A1312"
 <1..4417
 /gene="K1A1A1312"
 /note="start codon is not identified."
 /codon_start=2
 /product="K1A1A1312 protein"
 /protein_id="BA02550.1"
 /db_xref="GI:7242979"
 /translation="NSDHTVAVLSLCSGMIGTFRSHDGYFIEFLQSMDEDEDEECONKPHIIRRSAPDREPSTGRHACDTESEKRNHRSDKRRKARKMERINLAGDVAALNSGLATPFAVAGKTDPTREKTRHRTKRELSYPRFVAVLVADRRVYSGENLQHYILTMSVIAIYKIDPSIGLNIYVILVHNEDEGSIENAOITLKNCOHOMSKNSPGIHHDPVAVLTPRODIGHADKDPDGLAELGTTIDDPYRSGSISDSDGLSVAFTTAH EGVHHPMPHDNNKCKEVEVSRQNYMATTLPYTNPMWKSCKSKYITTEFLDIPYGECLNDEPSEKRYLVQLDGLVNLNKKOGLLTPGSOVCPTMQRRLMKNVAVH KGRTOHPMADGTECPKHKYKGFVPEMDVPLTDSWGSWSPFGCSHRFGGI KTAIRCNRPKNGKRYCVGRMRKFKSCNTEBQLKRDYFDEQCAHDFGDFNLING LLPNVWPKYSGLIMKDKRCLEGRVAGNAYQLDRVYLDGTPCODINDLCOVGLC RQAGCDVILNSKARRKCGVCGGDNSSCKTVAQTPVHNGVYVYVIRIPAGAVNDIV OHSFSGEPTDDNVTALSSKGEPLLNENFVMAKREIIRGNVAVVYSGSEPTAVAEIN STDRIEQLLOVLSVGRKLVNPDVRYSPNIPEDKROQRTVNSHGFWQACSKPQGER KRLLVGTRESQQLVSDQRCDRLPQPHILTEPCGTDLEKWHVHVSSECSAOCGLDYR TLDLYCAKYSRLDKTEKVDGDFGCSNHPKPSNRKCSGECNTEGMRYSAMTECSKCD GGTORRRAICVNRINDVLDSDKCTHOKKVTIQRCSEFPQWMSGDMSECLVTCGSGH KHRQWCOFGEFLNDRMCDPEKPTMOTCOOPCEASMQAGMGCCSYTCGGQVQLR AVKCIIGVMSVVDNDNCNATREPTDODELSECSHPRAAPTRRSYTSAPRTORF GSNTPCSATGCKGKTRMYVSCRDENSVADSESACTLPPRVAKEECSYVPCQWAKLD WSSCVTGGGSRATROYVCVNSDHYTDRSECDQITPEFDQCSNRPQRTPDSGL AOHFPONBDYRPRASPSRTHVLVGNQWRTGFWAGCSSTCAGSORVVVACDQENGYT ANDCVERIKRDEQACSGPCQWAAVGNWGECKLCCGGIRTRLVVCOQSNQERFEDL SCETLIDKPPRDEQCNTHACPHDAAMSTGPMSCSVSCGRHNRKORNVYCAKQSHLES DYCENLAKPHEHRCRGGRCRPMKAGAMSOCSVSCGQVQQRHVCQIGIHTKLTARTE CNPYTRPESHEDCGPRCPLTYTMAEBEMOECTKTCGEGSRVRYKVVQVCDNKNKNEVHAR CDVSKRPVDRSSCSIQPCFVWITGEMSEVPSMEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t
 ORIGIN
 AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803
 Initial Score = 7 Optimized Score = 42 Significance = -0.43
 Residue Identity = 19% Matches = 66 Mismatches = 149
 Gaps = 121 Conservative Substitutions = 0
 Translation Frame= 1
 X 10 20
 EMGSPDAAAVRR-DRLHRRQVX

FEATURES
 source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="fhl1767"
 /tissue="brain"
 /clone_lib="pbluescriptII SK plus"
 /gene="K1A1A1312"
 <1..4417
 /gene="K1A1A1312"
 /note="start codon is not identified."
 /codon_start=2
 /product="K1A1A1312 protein"
 /protein_id="BA02550.1"
 /db_xref="GI:7242979"
 /translation="NSDHTVAVLSLCSGMIGTFRSHDGYFIEFLQSMDEDEDEECONKPHIIRRSAPDREPSTGRHACDTESEKRNHRSDKRRKARKMERINLAGDVAALNSGLATPFAVAGKTDPTREKTRHRTKRELSYPRFVAVLVADRRVYSGENLQHYILTMSVIAIYKIDPSIGLNIYVILVHNEDEGSIENAOITLKNCOHOMSKNSPGIHHDPVAVLTPRODIGHADKDPDGLAELGTTIDDPYRSGSISDSDGLSVAFTTAH EGVHHPMPHDNNKCKEVEVSRQNYMATTLPYTNPMWKSCKSKYITTEFLDIPYGECLNDEPSEKRYLVQLDGLVNLNKKOGLLTPGSOVCPTMQRRLMKNVAVH KGRTOHPMADGTECPKHKYKGFVPEMDVPLTDSWGSWSPFGCSHRFGGI KTAIRCNRPKNGKRYCVGRMRKFKSCNTEBQLKRDYFDEQCAHDFGDFNLING LLPNVWPKYSGLIMKDKRCLEGRVAGNAYQLDRVYLDGTPCODINDLCOVGLC RQAGCDVILNSKARRKCGVCGGDNSSCKTVAQTPVHNGVYVYVIRIPAGAVNDIV OHSFSGEPTDDNVTALSSKGEPLLNENFVMAKREIIRGNVAVVYSGSEPTAVAEIN STDRIEQLLOVLSVGRKLVNPDVRYSPNIPEDKROQRTVNSHGFWQACSKPQGER KRLLVGTRESQQLVSDQRCDRLPQPHILTEPCGTDLEKWHVHVSSECSAOCGLDYR TLDLYCAKYSRLDKTEKVDGDFGCSNHPKPSNRKCSGECNTEGMRYSAMTECSKCD GGTORRRAICVNRINDVLDSDKCTHOKKVTIQRCSEFPQWMSGDMSECLVTCGSGH KHRQWCOFGEFLNDRMCDPEKPTMOTCOOPCEASMQAGMGCCSYTCGGQVQLR AVKCIIGVMSVVDNDNCNATREPTDODELSECSHPRAAPTRRSYTSAPRTORF GSNTPCSATGCKGKTRMYVSCRDENSVADSESACTLPPRVAKEECSYVPCQWAKLD WSSCVTGGGSRATROYVCVNSDHYTDRSECDQITPEFDQCSNRPQRTPDSGL AOHFPONBDYRPRASPSRTHVLVGNQWRTGFWAGCSSTCAGSORVVVACDQENGYT ANDCVERIKRDEQACSGPCQWAAVGNWGECKLCCGGIRTRLVVCOQSNQERFEDL SCETLIDKPPRDEQCNTHACPHDAAMSTGPMSCSVSCGRHNRKORNVYCAKQSHLES DYCENLAKPHEHRCRGGRCRPMKAGAMSOCSVSCGQVQQRHVCQIGIHTKLTARTE CNPYTRPESHEDCGPRCPLTYTMAEBEMOECTKTCGEGSRVRYKVVQVCDNKNKNEVHAR CDVSKRPVDRSSCSIQPCFVWITGEMSEVPSMEL"

OLRANGHROPRLRNKGIHPVSKWGLFYXTTTVYGYXTRRKRKQTKOTPHHIAQRPERKALMKNKCMXHLATPK
 10 20 30 40 50 60 70
 LLEITLSEYELIYS--PIRNNALGEPERPVIWHKFKTRRSI-----NSATDP--WPAPASS
 30 40 50 60 70
 XAQRKOE-ENQSKKMKGRKDXPR-----WKRSSIKQRLSNRGIETLIMXDDGQHKREDYQAKRTFFIL
 80 90 100 110 120 130
 SSSSTSSQAHVRLSAFGGQPLFNLTANAGFTAPLFTVTLTCTPCVN-----DTEKFSSEBAILKICFY
 140 150 160 170 180 190
 STVYCS-----LGGGRQD-----NGEIPMKRPTSLYFNENVNSIYLAKPKYKFFHYCYCEL-----
 140 150 160 170 180
 KGVVNTNSEHT--AVISLCSGMIGTFRSHDGYFIEFLQSMDEQ-----EDEEONKPHIILYRRSAPQ--
 140 150 160 170 180 190
 ---NDSXHTGMATHIXCS-----DNIKLLPVAALFEQSRMNSXYCCSINK--TGYLOSSRQMX
 190 200 210 220 230 240
 --REPSTGRHACDTESEKRNHRSDKRRKARKE--RINLAGDVAALNS-----GLATEAFSAVGN
 200 210 220 230 240 250
 YLRPGTGNH-----LKSJKLFLYXKXKMTLSEFYDPRXAGPCVQVXASXXQOQMXRRRREXPACGHS
 250 260 270 280 290 300
 KT-----DNTREKRTHRRRTKR X
 260 270 280 290 300
 NPELLHQRLDVLVYKXSKLYHXYVFRHMLRVAFXRPTXIQTLTPACPTARHNPLOREXHYKIDFWTRFSGVPIYD
 310 320 330 340 350 360 370
 AVQPLVALXQRKQSTGRPLRDSAHTELRMDGVRAWKALDQWIIICSGRNGCPRDRWILKLESLMILDMNMRGH
 380 390 400 410 420 430 440 450
 ONSHSHVQQTFRKWKWKIILCPTXNEIYVLOHAGMSQAEARLPRXVTVCSLXKREAFHQRSASQCALGFXIQWN
 460 470 480 490 500 510 520
 SDEGRVQVVLQSGRREHSLLSAROSRDRANSILMRPGHKYXLCGPRLPASMSRSCFKLKSREFXMMGIMWRXRFM
 530 540 550 560 570 580 590
 CCGRVQWVADCCRRKXKXNLSHAYARFFASGFVGGKVVQPRCTLPFQOYSNXXRXTSALLLEQSMAMASMQXVLP
 670 680 690 700 710 720 730
 GTEETKTLHQGIXSAYCFKSKWRSAPRMTYHXTLWYRLXRPVACCOEQEXMXCPWLDLPHIGHLICOLIXQA
 740 750 760 770 780 790 800 810
 GWEDXGXXXWFLQDQSPQFKQPKMLRGXKHGWLALFCLDXMFKKLLXMRDPEKGYLCOYPRKXTGXQOIMHTS
 820 830 840 850 860 870 880
 RESYHSEVQXVPLSTVEIWRKLVYLGHLMKRAXMADPGLVSWMKSIXKXNVXPRXDQANITYADLSAAGKTLA
 890 900 910 920 930 940 950
 GGSILGTQOCNHLMTGIPAKSSEHNDLVVSGRXQXKXCSNXTNXYPGIAXITIMSSSPSCGNEEKHITQCTKN
 960 970 980 990 1000 1010 1020
 PVALVWLDLMLSHLMEKRYDETRDLPKXEMICGXRECLCTYPAKXSGKGRMPCDTLAMEGIGLELLLCDLMA
 1030 1040 1050 1060 1070 1080 1090
 RXGNPASYDCOLOXPRRSEXXVXPLGRPNKXPLGLFNHYTPSKDRPRLWLSAFLPKXGLSSPEROPORPHPCAR
 1100 1110 1120 1130 1140 1150 1160 1170
 WKRVENWPRIGSMFOYLQWRIPRPAACCSMGSKYIHRKRLCGENKTKXXAKSLXIRPLDSSVGLMOLGRVHXAAWM
 1180 1190 1200 1210 1220 1230 1240
 RHRNKTGGLSAVQRXVSRFELXNSXHTSRSSXAVXHNGLSTRRCMEYWPBELVFCLLMRAAXHTTKCLLHGRK
 1250 1260 1270 1280 1290 1300

WKPKKLLXAPGXATW
1320 1330




```

ASR COUNT 1471 a 1112 c 1362 g 1194 t
GCIN
AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ...
Initial Score = 8 Optimized Score = 30 Significance = -0.39
Residue Identity = 19% Matches = 45 Mismatches = 112
Gaps = 79 Conservative Substitutions = 112
Translation Frame= 4

```

```

MFTTGLPLKRYXISIVHRKIYERKMLVLEILRLIKIIFLXSCLLNRHNTSGSKXGVCPLDCCFKTKNFSPH
10 20 30 40 50 60 70
FOYRKHIIIVQIESLXKIRSGNARPHFCATAGVKEIILLPSHMKCFVTVVAISPCCLEPCTTIFSYTXLSSG
80 90 100 110 120 130 140
TSQHPMSLQPLXXFLLYVANTRSXITTECRPEGTXYRPSLSLPLCIGIILAMESXGLVRRKHTVHSGSGGXPO
150 160 170 180 190 200 210
VSHIPALIPPLIPPLSGHGKXADDGYSSODGTSDDHSRPLYQIYSGCGKLDLSRSTGILLISORAPCTISLSS
220 230 240 250 260 270 280
THHTTLRYLEPSPOVLVHSGHSSALQVXRGRHGFWSGSRSDGLVYGLHVSVAIILCVXPIXOPTCLCCTPRPOD
290 300 310 320 330 340 350 360

```

```

TEHXLQAPAFHLGHLRPHRHLPCGLARCIQXSLSKWLPSPAMQXTPRCLCPRPEDEHELOGPVLHAASCG
370 380 390 400 410 420 X 430
10 20 30 40 50 60
EVLVYVADNR-----AVSYHGENL---OHYITLMSIVASISYKDPSTIGMLINIVLNLIVHNEODGP-
440 450 460 470 480
QACVLHCSSRSGSLRISQKSGNRSPLDKRXQTTSLVLMPPHSLVHSP-----QLPXAXHXGQ-GPD
510 520 530 540 550
-----SISNNAQTTLKKNFGOMQHSKNSPFGSIHHDV-----AVLLTFQDIDICRAH---DKCDPLGLAELG
560
SOALCSSGGFLST-----OSFAVYPPSSXHTTTRRMDPPAQVLEHARPOGVLHWPFPPTVWVRLGALRIG
130 140 150 160 170
TICDPRKSSISDSGLSTAFPTLHHLGIVFWMPPHDNKKCKEYGVKSPQH-----VWAPR
180 X
LNRYTTPMWMMSKC
-----KXSS--FWKGC-----XAKPLSGVLXGHGDMQMSWSVSGISXMSHSLRSITMSLXLTHITCRVALP
560 570 580 590 600 610

```

```

CPQVTEQELQSKAFHFCPOGVTEHSSSPATGIGRVAQALSSATEPESSRQTLVILRVPPQVAEHGVDENRHM
620 630 640 650 660 670 680
VLGALYVLLVSGAAGGGXHDGNSQSWSVGLVALDQSSISSTPTDIXVPMHFTALSWPCROVPLHFCQGRA
690 700 710 720 730 740 750
CODAHSGCXQVCIDVGLVSGSHILSEFNSSPPXHQVTRCCLCFPPQVTRHSDOSPDPFHCGGNSILHXKWTFS
760 770 780 790 800 810 820 830
KCVHLSSSTSPRVLTLQALLMWPSSOLFENHVOAEKRORPVLIHSPHNSRLLGLGWLKPKSPSTSVFSS
840 850 860 870 880 890 900
SLLYLAQXMSVNRKXKPHMALHSLDLATQHLRSQSVPOGQSVKCPGWSRSHLXSEVTVSXSDSLVQTSFRFS
910 920 930 940 950 960 970
PMQGLIHCXGPIFLFOXNCGLSISGILKEKXRTSGLYNFPDVKTKXSKSSGSMRSEYELIHTVAVSDPLYSTTA
980 990 1000 1010 1020 1030 1040
FPMRISLAIYVTKRPFSSKNSLILLDKAKXLSSSSVSPKLCRPSILVAPAGINTVYLXPKYCTVLANVPATVL
1050 1060 1070 1080 1090 1100 1110
HEELSPQTPHLILRPAFEFKTKXSHBACRQRPWTQISFVSWFQGVPSITLSRXXXXAAVPPATLQNNLHRSPFR
1120 1130 1140 1150 1160 1170 1180 1190

```

```

IDLYIGTORITGSRPLMKCPFSKXAHSSKRSRCLRHGSLVDLQNLFPPTQYFPFFPFGSGLHSRMAVL
1200 1210 1220 1230 1240 1250 1260
MPPHYLQVPRKGLQLPDPPSVTGTSTISLGTQNPYLQCFPQSGSHVSPNOGVXVROPICPLPRLTLHQSRLLHC
1270 1280 1290 1290 1300 1310 1320 1330
IYGHXTEBGRPKINSHCLFLLXRMBSGWTGKXGGLDSSLSKHSRXPYKRSKSYVYRFLHPDHQIGLYXKESV
1340 1350 1360 1370 1380 1390 1400
GAMFCWGLITPSSLHLLSSXGMLNTPSSWAIYKAVLNPILSSLEQLDIXGOSQWPPSARPKVSHLSXALQI
1410 1420 1430 1440 1450 1460 1470
SCLVFRFAVSKWIPRGLPPECCOHMQKFNVYVXALKDMEGSCSLXITTKFTITLILKPIILGSLXITATDI
1480 1490 1500 1510 1520 1530 1540 1550
KVKIKXKRPSPWYETIILLASATTKTSTNRNGXDKKREVLIMVLESLVLELXAMNASYAKRPLFNAATSPARL
1560 1570 1580 1590 1600 1610 1620
ILSPH

```

```

3. US-09-972-467-2 (289-478)
AB037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
TOIG of: ab037733 check: 4803 from: 1 to: 5139
LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
ACCESSION AB037733
VERSION AB037733.1 GI:7242978
KEYWORDS Homo sapiens brain cDNA to mRNA, clone_11b;PBLuescriptII SK plus
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (sites)
Nagase,T., Kikuno,R., Ishikawa,K.I., Hirotsawa,M. and Ohara,O.
Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
DNA Res. 7 (1), 65-73 (2000)
JOURNAL MEDLINE 2 (bases 1 to 5139)
REFERENCE 20181126
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba

```

292-0812, Japan (E-mail:cdhainfo@kazusa.or.jp, URL:http://www.kazusa.or.jp/hunge/, Tel:+81-438-52-3913, Fax:+81-438-52-3914)

FEATURES

source Location/Qualifiers
1.5139
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="fhl1767"
/tissue_type="brain"
/clone_lib="pbluescriptII SK plus"
1..4417
/gene="K1AA1312"
<!--.4417
/gene="K1AA1312"
/feature_start=2
/feature_end=2
/feature_note="Start codon is not identified."
/protein="K1AA1312 protein"
/db_xref="GI:7242979"
/db_xref="refseq:NM_012249.1"
/db_xref="trnsc:trnsc:UAG

BASE COUNT 1471 a 1112 c 1362 g 1194 t
AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803
Initial Score 8 Optimized Score 34 Significance = -0.39
Residue Identity 22 Matches 45 Mismatches = 128
Gaps 25 Conservative Substitutions = 0
Translation Frame 3
GSLNVPEFKKRTGASGSAEMOGTOPPTSEFTEEMXMMELVARTOMISYSRAFPACKLDAIMEYTXQKPEEIWGF
520 530 540 550 560 570 580
VVAIILHAKQWQHHLIQIYIMVTTILMSEFQLVLPRLIMGSGTVSOGKQTMTTXXLYQAVKVNCSXMMETLLSQWP
590 600 610 620 630 640 650
KGRFALGMLXKSTVGRPLPKKELTQOIALSKNFRCRCMESCTTPMAYALISTPDLKINSSFTTGYVMGNGK
660 670 680 690 700 710 720 730
HAVNPAGKNGENLFPAGNLISLFLKDAIGCPSTDLTLNPVQVVTXGCMLEPAGVNVVPSAVWYATAWTTS
740 750 760 770 780 790 800
TVFNIAQMGMLRRLIMVFPAAAIPIQVTKNAQGNVTRVAGALIPGILNQKAVATVGPREGGLVSTPEMMIYW
810 820 830 840 850 860 870
MTANAHIKRRLPFGAVSVLHSGNLETGOSAVSPEVKGISTARSGSLVKIDIXMIETCIVTLRPSSHLCRLVS
880 890 900 910 920 930 940

SRNVHPRGRVPGDASVSLVDRDTSXEOXNLSGLICQMXYMPTVMQDLIPRTVYVHHYILPDLRKRGE
950 960 970 980 990 1000
AHTVYQDEPSGDLGSLRPAQRDRLVGVKVPQKXDTSALEMRYALMLETTRVPVLRCDLQOMKRLKHPGNGRPMWGA
1020 1030 1040 1050 1060 1070 1080 1090

LALXPGYGVQSGKXVSTTVTTXSIGVSPRRIISOILTRTPCHNALKBQVVAALSTPSKMTIYVPGAPA
1100 1110 1120 1130 1140 1150
PAAKCSVETSGELAPRGNHVPVPLADPSSVLLVVRKMDPRPTVTRKXNLSKPEVNPALVLSGLMATGE
1170 1180 1190 1200 1210 1220
SALSVCVEAXHBDQMWVSSGPTVNGEQIAVXAVFLNPLVIVSVMMLVHTTLEHYLALGARVLSLVEGIVNNE
1240 1250 1260 1270 1280 1290 1300

MFLAQKQMEAIKXVITYVSTWLSHMQTESAEDEDAFNGKLAGVSALCPVAEAVSRGMVAARSEHTKXPRPS
1310 1320 1330 1340 1350 1360 1370
AHTHPRDRSRNATAKAHNVSPSTLGGORNGKAPRPAKAPGTAARWCVMHTYKTRCMGHANVYXASGRMIVKAVV
1380 1390 1400 1410 1420 1430 1440 1450

CNPASMSGSOENGQRTRRPGNCNHRQLSHGLVAVSRYGCVQVEGRIEATRPICHCGLVFXOVLSP-RIE
1460 1470 1480 1490 1500 1510 1520
VYVADNBNRMYSHGENLOHYLL- TLMSTIVAYKDPISGMLNIVILNIIHNEODGSPISFNAQTLLKN
1530 1540 1550 1560 1570 1580

FCQ- VONSKNSVGGIHNHTAVLVTGTR- DICRANHDKCDTLGLAEGLTICDPYRSCSISSEDGLSTAF- TIANE
80 90 100 110 120 130 140
PQQLLQKTSVCCVKG----AESLSTLQPLRGRNVVXRYQDTXPKFKKGGXGLVCELSXLENGXNMFYNYVIG
1590 1600 1610 1620 1630 1640 1650

LNHF- NMFHDNNKCKEKGVYK- SPOHVMAPTLNFYTPMMWMSKC
160 170 180 190
KQTPMNMNQKCACLNLSFRGRKXFSIISEVSOVPRF-----HISCAACLSRPFLMVINQXXS
1660 1670 1680 1690 X 1700 1710

4. US-09-972-467-2 (289-478)
ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
TOIG of: ab037733 check: 4803 from: 1 to: 5139
LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for K1AA1312 protein, partial cds.
ACCESSION AB037733
VERSION AB037733.1 GI:7242978
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiinda; Homo.
REFERENCE 1 (sites)
AUTHORS Nagase, T., Kikuno, R., Ishikawa, K. I., Hirose, M., and Ohara, O.
TITLE Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
JOURNAL DNA Res. 7 (1), 65-73 (2000)
MEDLINE 20181126
REFERENCE 2 (bases 1 to 5139)
AUTHORS Ohara, O., Nagase, T., and Kikuno, R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba

292-0812, Japan (E-mail: cdnaha@fokkazusa.or.jp, URL: http://www.kazusa.or.jp/huge/, Tel: +81-438-52-3913, Fax: +81-438-52-3914)

FEATURES source Location/Qualifiers 1..5139

gene /organism="Homo sapiens" /db_xref="taxon:9606" /clone="fhl1757" /issue_type="brain" /clone_lib="pbluscriptII sk plus" 1..4417

CDS <1..4417 /gene="K1AA1312" /note="Start codon is not identified." /codon_start=2 /product="K1AA1312 protein" /protein_id="BAA92550.1" /db_xref="GI:7242979"

ORIGIN 1112 c 1362 g 1194 t ...

AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ...

Final Score = 8 Optimized Score = 28 Significance = -0.39 ...

SASROSDPANNLSLPHGKHYLCSGRLPASMMRSCGFKLKSPERKMMGIMMXXFPQNSGRNIYYSTLMLDQYCGP ...

ORXKMLRGXNHNALFCIDXMFKAKLRRDDREKGYLCOYPRKXCTXOQMHRSREYHSVQVPLSTYEIFI ...

DETROLPRXEMICGXRECLCYRAXTSGKGRMCDTLMMMEGIGLELLCDLIMARKXGNPASDQCLOXPRDRS ...

IS-----FNAOTTLKLNFCQMHNSKSPGCIHNDTAVILFRDIDICRAHDKDITLGLAELGTICDPYRSCS ...

TRLPRTVSPHLEGGMARMHODLRRLOVYRQGVYCGYQOKRGMGTLXREDAAGPYKXKXATLRYGLD ...

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000 ...

KEYWORDS Homo sapiens brain cDNA to mRNA, clone_lib:pbluescriptII SK plus
SOURCE clone:fl11767.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS Nagase,T., Kikuno,R., Ishikawa,K.I., Hiroseawa,M. and Ohara,O.
TITLE 1 (sites)
JOURNAL Nagase,T., Kikuno,R., Ishikawa,K.I., Hiroseawa,M. and Ohara,O.
MEDLINE 2 (bases 1 to 5139)
REFERENCE Ohara,O., Nagase,T. and Kikuno,R.
AUTHORS Direct Submission
TITLE Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
JOURNAL Laboratory of DNA Technology, 1532-3 Yama, Kisarazu, Chiba
292-0812, Japan (E-mail:cdna@in.foe.kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
FEATURES
source Location/Organism
1..5139 /organism="Homo sapiens"
/DB_xref="taxon:9606"
/clone="fl11767"
/lisue_type="Brain"
/clone_lib="pbluescriptII SK plus"
1..4417
<1..4417
/gene="K1AA1312"
/feature="Start codon is not identified."
/codon_start=2
/product="K1AA1312 protein"
/protein_id="BAA92550.1"
/db_xref="GI:7242979"
/translation="NSERTAVISLCSGMIGTRSHDGYTEPFLDSMDQEDDEBQNK
PHIYIRRSAPQREPTGRHACTDSEKMKRHSKDKKTRARKWGERINLAGDVAALNSG
LATEAFSANGKTDTRREKTRHRTKRLPSYRPFVEVLAVANRWVSYHGELOHIL
TLMSIVASIKDPSIGNLINIVN.IVHNHODGSPISFNQTLTKRNCOMHSKNS
PGIHHDTAVLLTODICRAHDKDPLGLAELGTCDDYRSGSISEDSGLSFAPIAH
ELGAVNEMPHDDNNCKEGRYSQHWAPRINFEYTNMMSKCKRKYITTEPFIAT
ECLNPERSPRYPLVYDGLIILVYVKKCEILTFEGSSVCPYIMQCRKLMCNVAVH
KCCRTQHPFMAAGTECEPEGKHCYGFVCPKEMDVPVTVGSWGSPPFGTCSRTGGGI
KTAIRCNRPBENKNGKCYVGRMRKFKSCNPEKORDFEDECFAHFDCRHFNING
LIPNVMYKYSIGIIMKDRCKLFCVAGNTAVYLRDVIIDCTPCGDPTINDICVOGLC
ROAGCDHYLNSKARBDKGGVGGNSCKTVAGTFFNVAHYGVYVRIIPAGATNIDVR
OHSFSGEDDDKYLALSSKQEPILLNNGVYVMAKRETRIGNAVVEYSGSEFAVERIN
STDRIBQELDLVLSVGLKLNPDVRISENIPLDPRQDFYVMSHGWACSKPCQGER
KRLIVCTRESDDLTVSDQRDLRQPHITTEPCGTDCLRHMHVARSBSAQCGYR
TLIDICAKYSRLDGTKEVDGFCSSHRKSPNREKSGECNTGGRYSAMTECSKSD
GGTORRAICVNTVNDVLDSSKCTHOEKTIQRCSEPCQMKSGDMESECLVTEGGKH
KRWYVWCGEGRDLNDRMGDDEPTKPTSMOTQCPGCAQWAGCPCVTCGGCQYLR
AVKCIITGTMVYDNDSCAAATRPPTDQCELPSCPPAPAPETRRSTYSAPRDMFR
GSMTPGSAATCGKTRMRYVSCDENGSVADESACVTLRPAKREKSYVPCGMWALD
WSSCVYCGQGRATRQWVNVSDHVIDRSEDDOYIETDQDCSMPCQPTDPDSSL
AQHPFONEDYRPSASPSRTIHLVGNOMRTGFWAGSSTCAGGSRVRYVVDENGYT
ANDCVERIKIPDEORACESEGPOMWYGNMGWSSCTKCGGIRTRVLYVCCORRSGERDL
SCETLIDKPPDRROCNHACPHDAARSTGPMSSCSYCGRGRKORVWYVMSADPSHES
DYCKHLAKRPHGRKCRGRCRPRKAKGANSQCSYSGRQVQDQRHGGCGITKRLAREIE
CNPYTRPESERDCGPRCPVLYTWRAEMOECTKTCGSGESRYRKYVAVDNRNEVHGAR
CDVSKRPVDRSECSLQPCPEYVITWTEWSEVPSWEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t
ORIGIN
AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..
Initial Score = 7 Optimized score = 37 Significance = -0.40
Residue Identity = 23% Matches = 49 Mismatches = 137
Gaps = 25 Conservative Substitutions = 0
Translation Frame= 6

DYWFETIKKGLDKKCAQEDIKKNGTSDIIEYKFLPKLGVKQAOHWIMFKGVCLRLTFMYKXKQFQPF
10 20 30 40 50 60 70
X
LSYPRVEVLY-VADNRNVSTHGENLQHYIL-LMS
SSXELSHNSPDPKPLKNQWXXKXTPFLHGSWMSRSDSDAPPTOHVLCWCHSCLSLRPMHYNIPLYLTVFRH
80 90 100 110 120 130 140
IYAS-IYKDP-SIGMLNIVYVNLIVHNEODGSPIS--FNAQTT-----LKNPCOMHSKNSPGGIHHDT
150 160 170 180 190 200 210
LVAHYEFAPVLISSLCOXDIKVNHRMXXRDLRTIYTSALHTRNSLKLRTCKQTHS--SOWOI-GRV
150 160 170 180 190 200 210
100 110 120 130 140 150
AVLLRQDICRAHDKDPL-----GLAELGTCDDYRSGSISEDSGL-STAFTIHELGHVYVNM-
ASLPHSCTHPSTHPSATLRWLSYRWLQFPRGLXP-FSCDDPDLAQLQTAFTVNRPLAHVYACSPHILVE
220 230 240 250 260 270 280

160 170 180 190
NKCREEGVKSPQHWAPRLNFTYTNPMWMSKC
VVIHNNHLAVPQAFAGLGAFLPFLCPSVEGTPWALAVAFRLNSGVWVALGSGYFVCSDDLRAHMDLLYAS
290 300 310 X 320 330 340 350
ATGHRALPSPASEPFGASSSSALSVPMWLSOYLTVYTFXMASIFCPHAVNISLFPMPSTRTDRTRAPRASTPCS
360 370 380 390 400 410 420
VYWTSMCVTLTLTIGRPIKNTFQAIKRPVGVRLDHDQSCSYASTTQLSALSPVALISPLRKRAGTTGSLIRE
430 440 450 460 470 480 490 500
YSLHTVCGVSTFILLVNTPTLPGASATGTGTCSPGASPLVSTEHMGAAGAGRTIYLILEGVLSXATVWG
510 520 530 540 550 560 570
PRLAHHGCTVLSVFDIILVTLPIVDHVVTVVDVTHLDEGCPTLPPGHRARAPVQGLRPLRPGCHRTFFLCHMS
580 590 600 610 620 630 640
RQSGTGLTVSHRAILISADVYSHRPTPTSGXAMGPRPKSPDLSKCTVCAASPRFRSGMRTXHXFTVLGIS
650 660 670 680 690 700 710
WSSCCTIYVIVYHXAISRNDAPHCXKLVLSLSTPVALSPGTRRDEGCTFRLLTSLHRCMGLGRVHSHIIXSIF
720 730 740 750 760 770 780
TKLTPDLAVLMPFSTGDOALXPRVSRFPMTRELPARLNGNPLMCAFAVIOYIISGIDTNSPSPDLGPTVAF
790 800 810 820 830 840 850
XHPSPGRFAPRATRYTPFAFPLVAMFGMAAKTIIINLSDLPIDRALRGTVDVOCVAVQATIGTTFPAGWMP
870 880 890 900 910 920 930
PQVTVCTTGFSSVNSHLGQPIASLIRNSKLIIRFGANKRFPPLAGFTACLPMDVTVPKLLRFIFMWNIER
940 950 960 970 980 990 1000
IAYIVVQLSHRONIKQKFLMLNATIXVNSFYSGISGPIVLYHSIEMANRPFSGHCNKYSIXOEFPLTAXXSV
1010 1020 1030 1040 1050 1060 1070
VIVYCFPKETVLPKPHINISGWSNDSHSIVTIVMCSKCHCFAXRIIATVPTFISPGFYXNMIIASSLPAK
1080 1090 1100 1110 1120 1130 1140
ALDDITICVLAATRSSIYHSVSKLIVGVCYPCGSAEQLADVYLHONSTVFRDPAHIGKOTVDVYKMLPKVYKSTLEI
1150 1160 1170 1180 1190 1200 1210 1220
SFVSLLEFTWLRVAGLKEFHSSTSYVVFSTIIFMFWSVYALNSGCFDAPRSCSAGSGKGFPTSPGICGHDIHFFG
1230 1240 1250 1260 1270 1280 1290
NKSIIAVALSRLALRPIGPGCVL
1300 1310

6. US-09-972-467-2 (289-478)
ab037733 TOIG OF: ab037733 check: 4803 from: 1 to: 5139

TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
ACCESSION AB037733
VERSION AB037733.1 GI:7242978

ORGANISM Homo sapiens brain cDNA to mRNA, clone_lib:pbLuescriptII SK plus
Clone:fh1167.
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE Nagase,T., Kikuno,R., Ishikawa,K.I., Hirosewa,M. and Ohara,O.
Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
DNA Res. 7 (1), 65-73 (2000)
2 (bases 1 to 5139)
Ohara,O., Nagase,T. and Kikuno,R.
Direct Submission
Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:odnainfo@kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)

FEATURES
source
Location/Qualifiers
1..5139
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="fh1167"
/tissue_type="brain"
/clone_lib="pbLuescriptII SK plus"
1..4417
/gene="KIAA1312"
<1..4417
/gene="KIAA1312"
/feature_start="not identified."

gene
CDS
/odon_start=2
/product="KIAA1312 protein"
/protein_id="BA092550.1"
/translat="NSEHTAVVLSLCSGMLGTPRSHDQVTEPLQSMDEQDEEERONK
PHIYRSAPDRPSPGSRHACDTPSEHKHNSDKKTKTRKRGKERRINLAGVYAAIANSK
LATFAASAYGKNTDNTREKTRKTRKRPSTYRPEVLLVAVANRWVSTHGEWQHTIL
TLMSTIASIKRPSSTGNLNIYVNLVHNQDQSPISFNQTLKNEFCWQSHKNS
PGIHNHTAVLLRQDLCRAHDKCDPLGLAELGTICDPRRCSISEDSGLSTAFPIAH
EGLHVNEMPHDNNKCKEGLVSPQHVMAPIFNPTNPMWMSKSRKYITFELDTGYG
ECLINRESRYPPLVYOLPGLIYLVNKOCELIFFGSGVCPYPMGCRRLMKNVNVH
KGRQDHTPMADGETEGEPGKHCYGFVCKEMDVLTGSGMSWSPFGTCSRGGGI
KTAIRECNREPEKNGKCYVGRMKKSCNTEPCLVKOKRDRPFDQCAHFDKHPNNG
LIPNVWVAVYSGIILMKDKCKLFCVAVAGTAYOJLDYIDSTPGGQDITNDICVQGLC
RQAGDHYNSKARADKCGVCGGDNSSCKTAVTNYHYGNVYRIPAGVINTDVR
OHSFSEITDDNVLALSSKGEFLNNGNVTVMARERLIGNAVVEYSSEFAVYRIN
SVDRIEQLLELLQVLSGKLDYNDVRSFNIPEDKPOQFPMOASPCOGER
KRLVCTRESQDQVSDKCDRLPQGHTEFCGTDCLRMHVASRSESAOCCGGYR
TLDICAKYSLRDGTEKYDGFCSHSPKSRKSGEENGGRRYSAMTEKSCG
GGTORRAICVNTNRNDVLDKCTHQBKVTIQRGSEFPDQPKSGDMSKCYTTCGKH
KHROYWCQGEEDRLNDRMCDPEKPTSMQTCOQPCASQWQPMGQCVSTGQGLQR
AYKCIITGTVASVDDNDCAATRPDTIDCELPKPPPAVETRRSTYSARITQWRP
GSMTPCSATCGKGTMRMYVSCDENGSAVDEACATLPRPAKETSRTYSARITQWRP
WSSCVYCGQGRATQVQVNYSDHVIDRSECDQYIETDQDCSMSPQSTPGSLG
AQPPEQNEDEYRPSASPTHTVYLGNNOMRTGPMWACSSYCAAGSRRVYVODEDGYT
ANDQVRIKIPDEORACEGSPCOMAYVGNMVECTKLGCGGIRTRIVYVCCQNSGRRPDI
SCGELIDKPRDROCCMHACPHDAVSTGPMSSCSYCGRGRKORVYVMAQDGSILS
DYCKHLAKRHGRKRGRCRPMKAGANSQCSVSGCGRQVHVCQQLGTHIAETE
CNPITRPPVRESQDGRCPYLWRAEWEQECTKTCGEGSSRYRKYVCCVDDNKNREVGAR
CDVSKRPPVRESQSLQPCBYVWITGEMSEVSPWEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t
ORIGIN
AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803

Initial Score = 6 Optimized score = 33 Significance = -0.42
Residue Identity = 21% Matches = 47 Mismatches = 125
Gaps = 45 Conservative Substitutions = 0
Translation Frame= 5

XLLLVYVYHKKKGRKXALCTGRMKWYLRVYGGYRKLFFSSSEAVVXGTGTHLHVERGRLPAPYVNLKPKIASPI
10 20 30 40 50 60 70

X LSY-PRF-----VEVLYVADNRNMYSTYGENIQLHTLMLSTVASTYKDPG---IGMLIVTVN-L
60 70 80 90 100 110 120 130 140 150 160 170 180 190 200

EKLGNLTYXSRFLASLEKSGLVYLD--HISAMPQLEKRRCSILHACASLSQPLSHLVSLALQYFLINCL
80 90 100 110 120 130 140 150 160 170 180 190 200

IYVHNQDQDPSISFNFNQTLTKNFCWQH---SKNPFQIHNHTPAVLLRQDICRAHDKCDITGLAE-----
150 160 170 180 190 200

QAPHSILXVCSPCNN----FFFMPIGHQESQNVLDLKLGRD--HLL--HLSRFAVYKFEFLGKRAEDLSEMT
150 160 170 180 190 200

-----LGTICDPRYSCSISDPSGLSTFAFTTAHLELGHVFNMPHDDNKKCKEYKSPQHVMAPIPLNRYTP
210 220 230 240 250 260 270

OFFVAVNRAAGSLKSPFLHS--SLHSSLSCHSQAMA-EL---TWVTVPRVPLTILLXSRHTRRVAVYSPHGP
210 220 230 240 250 260 270

RMWMSK
190

AACSRISVYHARPRFCCHPTPPCGTWSLRRRSWCLIALPLPESKCRGDTVGLGSRVPRVYVWCMGCTRSIMLFC
X 280 290 300 310 320 330 340

VFRSDPRHSAVAALGHRGTQSTDSKROLSTWIGTIFLLGTETCAHVAHXEAVSNHPLNGFHLIPCSKHFVYVALDH
350 360 370 380 390 400 410

KRQNTSSKQOYMSQWRVYKNCVYTAHREVTYDEHSSNLETVYKRTADRPVYELCLNHTAXCTIPSCHKRP
420 430 440 450 460 470 480

EDKRIHRLRFAQVLEFSRHSRLRCIHFHEDIQOHAAGIRQHRVYVNMMLPRGQFSTGFHRAHCGMGWRSGDSD
490 500 510 520 530 540 550

PHRGRGAELSHLQGSFEGMVTWNSGQFLGVYRPGHTHDSRSGHCSXHTSLAGLRYLAAHRSQKSSSPRST
560 570 580 590 600 610 620 630

AHRVSONLPLRPLVAXGXHRHSRODQSHSLGSKXRISSGYLSHKWLSMGSKTQIATGFLVHMCQFSSPFCOL
640 650 660 670 680 690 700

GEDDMVYIHSFGYOLXLLHNYHNCPLRPLTKSQXCISLILLAGIPIVHKXKCVYRDRPARMIIPAADKSAKML
710 720 730 740 750 760 770

ANSQGTTFYHLLDLDHQTDRRPGGAYALHHRXPSYTLTSLQISTVDKGTHTCSEMKXISLDVICCCHRVVHFGYX
780 790 800 810 820 830 840

HXXPSSGSHRHSFLNIOSRONASAHPCYIPLSIFRHCGLVMDGCCKNHNSPSSHPACUYIMHSRCPMGSP
850 860 870 880 890 900 910

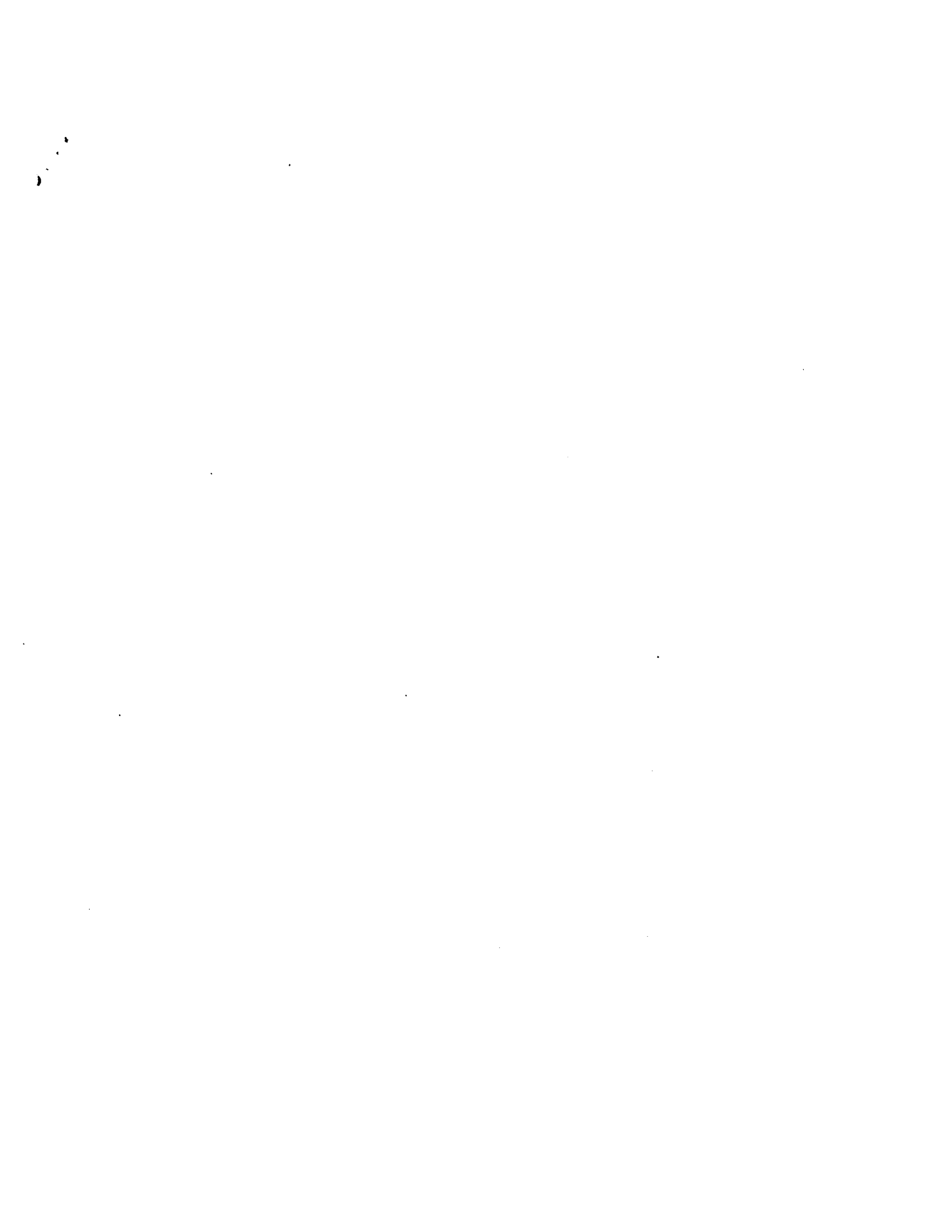
SHPTGNYIHSQWNAHSGHSLYRVOXCVQAGAADPIEQOKXADQIPKCKAVYVSPVLRGVYVCMIAAHNDCS
920 930 940 950 960 970 980 990

SKTAEVYLDLEVYKXNSVHRKCTTFRPTPRPEAKVLAQCDLLSXFFLRQSRGTCTLPHNSOCSEFPWPLXQGSF
1000 1010 1020 1030 1040 1050 1060

HLARLHYKLLKLSGCHYRFLRNCAAHQVXHNQLEGGROYCANNVYLYMPLDFPKKKNYRHNKRPHTIYLSG
1070 1080 1090 1100 1110 1120 1130

LLSLKNRDIQLAGKGRGHRNYLCPGSHKFEHNLISLCEADSRILCSPLCTTCTGSPSEFFNIXGSSAVHADR
1140 1150 1160 1170 1180 1190 1200

XGXNARSQSEHTVHLSGLASADMAPCCRGTXYISFVYVLSITFNHPLVVLVCSLTEMFLPKCPILMFPMSRRQRD
1210 1220 1230 1240 1250 1260 1270



KRKLVCTRESQDLTVSPQRCDDRLPQEGHTTEPQSTDDLRMHVVASRECSAOCGLIYR
 TLDITCAKYSRLDKTEKRVLDVDFGFCSSHPKSPNSRKGCGEENCTGMRWISATFESKSCD
 GGQRRRACIVNTRNDVLDJDSKCTHNEKVTIQRSEPCPCOMSGMDSFELVTCGKH
 KHRVQCOFGEDRLNDRMCDPEPKTSMQCOQECASMSQAGWQCCSVYQGOYQUR
 AKVIGITVMSVVDNDCMAATREPTDPODELPSCHPPEAPAREKSTYSAPRTQMF
 GAMPKSNATCGKTRMRYVSGRDNESVADBSACATLPPRYAKREKESVTPGQMKALD
 WSSCVTGGQGRATROYKCVNYSBHYVDRSECDDTYIPENDQCSMSRPOORPDSGL
 AOHPPQNDYRPRASASPRHVLGSGNOMRTRGPMWAGCSTACGSSQRRVYVQCOENGT
 ANDCVERIKRPEORACESGPHDAAMSTGPMWSSCSGGRHKONRYVCMKADGSHLES
 SCEILDKPRDEQCNTHACRPHDAAMSTGPMWSSCSGGRHKONRYVCMKADGSHLES
 DYCKHLAKPHGRKRCGRGRCPKKAKAGMSQCSVSCGVOORHVGCCOIGTHKTRQMF
 CNPYTRPESERDDCGRCRCLYTWRAPEMOBECTTKCGGSRVRYKVVYCDMKNEVHGAR
 CDVSKRPRDRESCSLDQPCREYVWITGEMSEVSWML"

BASE COUNT 1471 a 1112 c 1362 g 1194 t
 ORIGIN
 AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803
 Initial Score = 1471 Optimized score = 1471 Significance = 2.04
 Residue Identity = 1008 Matches = 1471 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0
 Translation Frame = 2

MGVSWATLTLTLVLDLAEWGSPPDAAAARVKKORLHPROYKLLLETUSEYDIYSPYRVALGSEDFPTNVHFKRT
 10 20 30 40 50 60 70
 RRINSATDMPWPAFASSSSSTSSQAHYRLSAFGQDFLENLNANMAGFIAPLFTVTLGTPGVYVQNTFYSEEE
 80 90 100 110 120 130 140
 150 160 170 180 190 200
 AELKHCHYKGYVYVNTNSHHRTAVISLCSGMLGTFRSHDQDFIETPLQSMDEQDEEONKPHIILYRRAPQREP
 NSEHTAVISLCSGMLGTFRSHDQDFIETPLQSMDEQDEEONKPHIILYRRAPQREP
 X 10 20 30 40 50
 220 230 240 250 260 270 280
 STGHACDSTSEHNKRHSKDKKTRRARKWGERINLADVAALNSGLATEFSAAGNKTDTREKTRHRRKTRF
 STGHNACDSTSEHNKRHSKDKKTRRARKWGERINLADVAALNSGLATEFSAAGNKTDTREKTRHRRKTRF
 60 70 80 90 100 110 120 130
 290 300 310 320 330 340 350 360
 LSYRFEVYLVVADNRKMSYHNGENLQHYIITLMSYVASYKDPISIGNLNIYVNLIVTJHNEDDGSPISFNA
 LSYRFEVYLVVADNRKMSYHNGENLQHYIITLMSYVASYKDPISIGNLNIYVNLIVTJHNEDDGSPISFNA
 140 150 160 170 180 190 200
 370 380 390 400 410 420 430
 QTLTKNFCSQOHNSKNSPFGIHNHDTAVLLTRQDICRAHDKCDLGLAELGTDYRSCSISSEDSGLSTAFIT
 OTTLKNFCSQOHNSKNSPFGIHNHDTAVLLTRQDICRAHDKCDLGLAELGTDYRSCSISSEDSGLSTAFIT
 210 220 230 240 250 260 270
 440 450 460 470 480 490 500
 AHELGHVFNMRHDDNNKCKEKEGVYKSPQHVWAPLNTFTNPMWMSKSKSRKITTPELDTGYECLLNRPESRY
 AHELGHVFNMRHDDNNKCKEKEGVYKSPQHVWAPLNTFTNPMWMSKSKSRKITTPELDTGYECLLNRPESRY
 280 290 300 310 320 330 340
 AHELGHVFNMRHDDNNKCKEKEGVYKSPQHVWAPLNTFTNPMWMSKSKSRKITTPELDTGYECLLNRPESRY
 510 520 530 540 550 560 570
 PLRYVQLRGITLNVKQKCELIFFPGSSQVCPYVMQCRRIKMCNNVNGVHNGGCRKQNTQTPPADGTECEPKHCKYGF
 PLRYVQLRGITLNVKQKCELIFFPGSSQVCPYVMQCRRIKMCNNVNGVHNGGCRKQNTQTPPADGTECEPKHCKYGF
 350 360 370 380 390 400 410
 580 590 600 610 620 630 640
 CVRPEMDVPLVTDGSMGSMSPGCTGSRKCGGGIKTALTECNRRPEPKNGKGYCVGRMKRCKSCNTEPCLKOKRD
 CVRPEMDVPLVTDGSMGSMSPGCTGSRKCGGGIKTALTECNRRPEPKNGKGYCVGRMKRCKSCNTEPCLKOKRD
 420 430 440 450 460 470 480 490
 650 660 670 680 690 700 710 720
 FRDEQCAHFDGCKHENNINGLPLNVKMPKYSGITLMDKCKLFCRVAAGNTAYYQLDRKIDYIDGTPCGGQDINDICV

FRDEQCAHFDGCKHENNINGLPLNVKMPKYSGITLMDKCKLFCRVAAGNTAYYQLDRKIDYIDGTPCGGQDINDICV
 500 510 520 530 540 550 560
 730 740 750 760 770 780 790
 800 810 820 830 840 850 860
 870 880 890 900 910 920 930
 940 950 960 970 980 990 1000
 1010 1020 1030 1040 1050 1060 1070 1080
 1090 1100 1110 1120 1130 1140 1150
 1160 1170 1180 1190 1200 1210 1220
 1230 1240 1250 1260 1270 1280 1290
 1300 1310 1320 1330 1340 1350 1360
 1370 1380 1390 1400 1410 1420 1430 1440
 1450 1460 1470 1480 1490 1500 1510
 1520 1530 1540 1550 1560 1570 1580
 1590 1600 1610 1620 1630 1640 1650
 1660 1670 1680 1690 1700 1710 1720
 1730 1740 1750 1760 1770 1780 1790
 1800 1810 1820 1830 1840 1850 1860
 1870 1880 1890 1900 1910 1920 1930
 1940 1950 1960 1970 1980 1990 2000
 2010 2020 2030 2040 2050 2060 2070 2080
 2090 2100 2110 2120 2130 2140 2150
 2160 2170 2180 2190 2200 2210 2220
 2230 2240 2250 2260 2270 2280 2290
 2300 2310 2320 2330 2340 2350 2360
 2370 2380 2390 2400 2410 2420 2430
 2440 2450 2460 2470 2480 2490 2500
 2510 2520 2530 2540 2550 2560 2570
 2580 2590 2600 2610 2620 2630 2640
 2650 2660 2670 2680 2690 2700 2710 2720
 2730 2740 2750 2760 2770 2780 2790
 2800 2810 2820 2830 2840 2850 2860
 2870 2880 2890 2900 2910 2920 2930
 2940 2950 2960 2970 2980 2990 3000
 3010 3020 3030 3040 3050 3060 3070 3080
 3090 3100 3110 3120 3130 3140 3150
 3160 3170 3180 3190 3200 3210 3220
 3230 3240 3250 3260 3270 3280 3290
 3300 3310 3320 3330 3340 3350 3360
 3370 3380 3390 3400 3410 3420 3430
 3440 3450 3460 3470 3480 3490 3500
 3510 3520 3530 3540 3550 3560 3570
 3580 3590 3600 3610 3620 3630 3640
 3650 3660 3670 3680 3690 3700 3710 3720
 3730 3740 3750 3760 3770 3780 3790
 3800 3810 3820 3830 3840 3850 3860
 3870 3880 3890 3900 3910 3920 3930
 3940 3950 3960 3970 3980 3990 4000
 4010 4020 4030 4040 4050 4060 4070 4080
 4090 4100 4110 4120 4130 4140 4150
 4160 4170 4180 4190 4200 4210 4220
 4230 4240 4250 4260 4270 4280 4290
 4300 4310 4320 4330 4340 4350 4360
 4370 4380 4390 4400 4410 4420 4430 4440
 4450 4460 4470 4480 4490 4500 4510
 4520 4530 4540 4550 4560 4570 4580
 4590 4600 4610 4620 4630 4640 4650
 4660 4670 4680 4690 4700 4710 4720
 4730 4740 4750 4760 4770 4780 4790
 4800 4810 4820 4830 4840 4850 4860
 4870 4880 4890 4900 4910 4920 4930
 4940 4950 4960 4970 4980 4990 5000
 5010 5020 5030 5040 5050 5060 5070 5080
 5090 5100 5110 5120 5130 5140 5150
 5160 5170 5180 5190 5200 5210 5220
 5230 5240 5250 5260 5270 5280 5290
 5300 5310 5320 5330 5340 5350 5360
 5370 5380 5390 5400 5410 5420 5430
 5440 5450 5460 5470 5480 5490 5500
 5510 5520 5530 5540 5550 5560 5570
 5580 5590 5600 5610 5620 5630 5640
 5650 5660 5670 5680 5690 5700 5710 5720
 5730 5740 5750 5760 5770 5780 5790
 5800 5810 5820 5830 5840 5850 5860
 5870 5880 5890 5900 5910 5920 5930
 5940 5950 5960 5970 5980 5990 6000
 6010 6020 6030 6040 6050 6060 6070 6080
 6090 6100 6110 6120 6130 6140 6150
 6160 6170 6180 6190 6200 6210 6220
 6230 6240 6250 6260 6270 6280 6290
 6300 6310 6320 6330 6340 6350 6360
 6370 6380 6390 6400 6410 6420 6430
 6440 6450 6460 6470 6480 6490 6500
 6510 6520 6530 6540 6550 6560 6570
 6580 6590 6600 6610 6620 6630 6640
 6650 6660 6670 6680 6690 6700 6710 6720
 6730 6740 6750 6760 6770 6780 6790
 6800 6810 6820 6830 6840 6850 6860
 6870 6880 6890 6900 6910 6920 6930
 6940 6950 6960 6970 6980 6990 7000
 7010 7020 7030 7040 7050 7060 7070 7080
 7090 7100 7110 7120 7130 7140 7150
 7160 7170 7180 7190 7200 7210 7220
 7230 7240 7250 7260 7270 7280 7290
 7300 7310 7320 7330 7340 7350 7360
 7370 7380 7390 7400 7410 7420 7430
 7440 7450 7460 7470 7480 7490 7500
 7510 7520 7530 7540 7550 7560 7570
 7580 7590 7600 7610 7620 7630 7640
 7650 7660 7670 7680 7690 7700 7710 7720
 7730 7740 7750 7760 7770 7780 7790
 7800 7810 7820 7830 7840 7850 7860
 7870 7880 7890 7900 7910 7920 7930
 7940 7950 7960 7970 7980 7990 8000
 8010 8020 8030 8040 8050 8060 8070 8080
 8090 8100 8110 8120 8130 8140 8150
 8160 8170 8180 8190 8200 8210 8220
 8230 8240 8250 8260 8270 8280 8290
 8300 8310 8320 8330 8340 8350 8360
 8370 8380 8390 8400 8410 8420 8430
 8440 8450 8460 8470 8480 8490 8500
 8510 8520 8530 8540 8550 8560 8570
 8580 8590 8600 8610 8620 8630 8640
 8650 8660 8670 8680 8690 8700 8710 8720
 8730 8740 8750 8760 8770 8780 8790
 8800 8810 8820 8830 8840 8850 8860
 8870 8880 8890 8900 8910 8920 8930
 8940 8950 8960 8970 8980 8990 9000
 9010 9020 9030 9040 9050 9060 9070 9080
 9090 9100 9110 9120 9130 9140 9150
 9160 9170 9180 9190 9200 9210 9220
 9230 9240 9250 9260 9270 9280 9290
 9300 9310 9320 9330 9340 9350 9360
 9370 9380 9390 9400 9410 9420 9430
 9440 9450 9460 9470 9480 9490 9500
 9510 9520 9530 9540 9550 9560 9570
 9580 9590 9600 9610 9620 9630 9640
 9650 9660 9670 9680 9690 9700 9710 9720
 9730 9740 9750 9760 9770 9780 9790
 9800 9810 9820 9830 9840 9850 9860
 9870 9880 9890 9900 9910 9920 9930
 9940 9950 9960 9970 9980 9990 10000


```

VOYVLOS -GREH-----SILSASROSDRWNMLPGRHXLYLCPGLPLPASWMSRSCFKLSPERXMMWGLMRRKXPFM
530      540      550      560      570      580      590
440      450      460      470      480      490
HELGHFNPNRHDNNKCKEKEGVKSPQHVAAPLNTYTPNMMWSKSKRYITEFL-----DTYGGEGLLNE
ONSGRNIYXSTLMLQYCGPNSWCYQYXCAAAQFLRGNRRKXOLLSFIIKQXRYLAKMKLCCNHGKQNSHWE
600      610      620      630      640      650      660
500      510      520      530      540      550      560
PESRRPLRYVLEFGLIYNVY-----KCELLIFGPGSVCYPMQCR-RLMGNVNVGHHGCRTPQHTPMA----
CCGR-----VQWVRDCRRKNXILNRSHXARTEFASG-FVGGKVVQPRCTLEFPQYSNKRXYSAVILLBQSWAMAS
670      680      690      700      710      720      730
-----DGTBEPKRN--CKYGFVYKEMDVRYLTDGSGWMSDPGTCSPRTCGGGIKTAIREC---NRPEKNG
OXLPRGTETKTCGLHGGIXSAYCFKSKM-----RSAPAPMTHYXLYLRYLAPREVACCOEQEMXKCFVWL
740      750      760      770      780      790
630      640      650      660      670      680      690
GKXCVGRMKFKSCNNEPRLCKOKRDRDEQCAHFEDGKHFNINGLLPLNVMWP-KYSGIIMKDRCKLFC----
GLPHIG-----HLLCQ-----IYQAGMEDXEGYXWFLQDQSPQTKORPKMLRGMXHGMLALFLCLDXM
800      700      710      720      730      740
-----RVAGNTAYUOLDRVIDGTSPCGQ-----DTNDICVQ-----GLCRQAGDHYLNSKARADKC
FKKLXMRDPREKGV-LGQYKXCTGXKQXQOHTSRESYHSEYQXVPLSTVEIWRVLVVLG--NIMKRAKXAPRG
860      870      880      890      900      910
750      760      770      780      790      800      810
GVCCGDMS-SCKTYVAGFNTVHY--GYNTVYVRIYPAGATINIDYQHSFSETDDNDNLALSSSGEFLLMGNF
920      930      940      950      960      970      980
LVSWXKRSIKXXNXYRYDQANITADLSAAGMCIAGGSLGTVOCHMTG-----IPASSENHMDL-----
VVTMAKREIRIGNAVVERYSG-SETAVYRINSTDRIBOELLQVLSVKLNPVRYRSPNPIEDKQDQGFWM
990      1000      1010      1020      1030
YVSGRXOXLKXCNSNTXYPGLXITIMSSSPSCPGNE-EKHIIQ-----CTKNP-----VAITWILD-----PML
820      830      840      850      860      870      880
SHGWRQACSKRCQGERKRLVCTPRE--SDQLYVSDR--CDRL-PQPGHITPECCGDCDLR-MHYAS-----
1040      1050      1060      1070      1080      1090      1100
SHLWERYPDETRQLPRXEWMLCGXRECLCYRAXYSGKGRMFCDDIMAMEGLGLELLCDDLMAKXGNPASYCQ
-----PSECGAQCGLGRTLDIYCAKYSRLDKT--EKVDDGFCSSHPKPR---SNREKCGSEGCNTGCGKRY
1110      1120      1130      1140      1150      1160      1170
LOXPRDRSE-VXKPEL-----YPRXKPGIFHVYTNMSKDRPOMLSSAPRLPKKLSSEPEPQPPHPCARPKP
1010      1020      1030      1040      1050      1060      1070
SAMMFCSKSCDGGQORRAICVUNTRNVDLSDSKCTHNEKVTIQRCSSEFPGOMKSGWSECLV--TCGKGNK
VBNPFLGSMFOYLWRIYPAACCCSGKWKIHRKRLCGENKTYXAKSLIXIRLSSVGLMQLGGRVYXAAVWMBHK
1180      1190      1200      1210      1220      1230      1240
1080      1090      1100      1110      1120      1130      1140
HROWWQCQGEDRLNDPDKPTSMQTCQAPPCAS--WQAGPMGOCOSVTCGGQYDLRVAKVC-IITSTYMSV
NKTGGLSAVQXXYYSREFLXNSXYTSRKAAXVYHHCCLSTRRCMEYWPLELVFCCLIMSRAAXYTTKCLLHKRMPK
1250      1260      1270      1280      1290      1300      1310
1150      1160      1170      1180      1190      1200
VDDNDCNNAATRPRTDQDELPSCHPAPAPETRRSTYSAPRTQWRFGS-----WTPCSA--TCGGGTMRMYYS
FRKXLLXAPGKAT-----NAQXVPRR--KMPQMESWRLESVLCVLMPPRRRAEAKCGJSDRNTQNS

```

```

1320      1330      1340      1350      1360      1370
1210      1220      1230      1240      1250      1260      1270
CRDNGSVADSEACATLPRRYAKKECSVTP--CGOWKA----LDMSSCSVTCGGRAFTRQYCVNISHVI
1380      1390      1400      1410      1420      1430      1440
QRDVOVQIHQDGVSTRLPFRP-----TVSPHLHEGRGMAMHODLRRRLQVPOGGVCGXQOKRGAMGLXRE
1280      1290      1300      1310      1320      1330      1340
DRSECDODYIPEYTDQDCMSMSPCRPTDPSG--LAQHFQNEDETRPFSASPSRTYLDLGNQMR--TGPWGCAC
QAAQKPKLXFATLRYVCLDHRRMVYRGVLTGVTIVSSAMAXXEQORDREKRNVDGLRLPARFATVNCVFS
1450      1460      1470      1480      1490      1500      1510
1350      1360      1370      1380      1390      1400
SSTCAGSGQRNVVVCQDENGNTTANDCYERTK-----PDEQRACESSGCRPQWAXGNNGECTKLCGGGTRPR-L
DKSSAFPS-----XNSLYA--KRECRKRWLSLSPFRSTFCSPYCP-----IGIKKLLDGLQLOTRHML
1520      1530      1540      1550      1560
1410      1420      1430      1440      1450      1460
VYCCQNGER-----FPDLSCCELLDKRPPREOCNTH-ACPHDAMS--TGPWSSCSVSCGRGKORNV
XGAKRQLSIRRYCCSARDYTRRDGNCDKLAHAYXREQLMFISSCGHALMMSVYRPPDS-----REAKMLDX
1570      1580      1590      1600      1610      1620
1470      1480      1490      1500      1510
KCMADGSHLSDYCKIHLAKPHGHRKCRGRCPRKMKAGAMSGCS--VSCGRGVOQ-----RHVGCQDGT
Y-----VREPNLKMGAELFG-----FKTIGANKRPLSXTGCVVYVXGTASBEENFLXYPKYLYK
1640      1650      1660      1670      1680
1530      1540      1550      1560      1570      1580      1590      1600
HKIARETECNPEYTRPSESRDQCGPRPLYTWRAREMOCRTYTCGSEGRSRYRKRKCVDDNKNVHGARGCDVSKR
HPEITLPPVHNAYLDPFKWXXTSSNH
1690      1700      1710      1720
1610      1620
PVDRESCSLQPCERYWITGEMSEVPSWEL
3. US-09-972-467-2 (1-1629)
AB037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
TOIG of: ab037733 check: 4803 from: 1 to: 5139
LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
ACCESSION AB037733
VERSION AB037733.1 GI:7242978
KEYWORDS
SOURCE Homo sapiens brain cDNA to mRNA, clone_11b:pbJunescriptII SK plus
ORGANISM Homo sapiens
Clone: fh1167.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (sites) Nagase,T., Kikuno,R., Ishikawa,K.I., Hirosewa,M. and Ohara,O.
Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
DNA Res. 7 (1), 65-73 (2000)
JOURNAL
20181126
MEDLINE
2 (bases 1 to 5139)
REFERENCE
Ohara,O., Nagase,T. and Kikuno,R.
Direct Submission
Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail: cdna.linfo@kazusa.or.jp,
URL: http://www.kazusa.or.jp/huge/, Tel: +81-438-52-3913,
Fax: +81-438-52-3914)
FEATURES
Location/Qualifiers

```

source 1..5139
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="fhl11767"
 /issue_type="Brain"
 /clone_libs="pbunescriptII sk plus"
 1..4417
 /gene="K1AA1312"
 <1..4417
 /gene="K1AA1312"
 /note="Start codon is not identified."
 /codon_start=2
 /product="K1AA1312 protein"
 /protein_id="BAA92550.1"
 /db_xref="GI:7242979"

AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..

Initial Score = 8 Optimized Score = 228 Significance = 0.41
 Residue Identity = 18% Matches = 337 Mismatches = 1069
 Gaps = 5 Conservative Substitutions = 0
 Translation Frame = 5

MOFVSWATLTLTLLVLDLAEEMGSPDAAAVRKRDLHPROYKLELLETSEYEIYSPINVAALGEPFPPNHHFKRT
 RRSINSATPWPAPASS---SSSTSSOAHNR-LSA-----FGQOFLPNTLANAGFLAPLFTVTLTG
 KYYLRYFGYRKLFFSSSEAVXCTGTHLHGERGLFAPYVNLKPKISADPFLKGNLTLYYSRFLASLEKSGL--
 30 40 50 60 70 80 90
 TPGVNOFKTYSSEBAELKRCFKYKGYVNTNSETTAIVISLCSGMGLGFRSHDDGDFTEPLQSDDEDEDEBQNK
 ---VLDHISANPQLEKX-----RRCSSLHTACASLSQLEPSSLVVSIALQYF---LILNCLQ-----A
 100 110 120 130 140
 PHIIYRRAKSP-----QRE--PSTGRACDTSEHKNR-HSKDKKTKTARKWGF-----RNLNLG---
 210 220 230 240 250
 PHSILXVCSPCNNFFMPLIGHGSEQNVDLKGLRDLHLHLSRFAYKEFLXGLKAEDELSENTOFTVANRAGSLK

150 160 170 180 190 200 210
 260 270 280 290 300 310 320
 DVAAINSL-ATEAFSAYGKNTDTRKTRHRTKRFSLSPREVEVLVAADRMSYHG--ENLQHYITL
 SPTFLHSHLSLCHSQMABELTWT-VPRFVPLTLLX-----RHTRRVA---NYSFHPRAASRHSVPH
 220 230 240 250 260 270 280
 330 340 350 360 370 380 390
 MSIYASIKDPDSIGNLIVIVNLIVIHNEBDDGPSISFNQPTLLKNCQWQSHKNSFGIHNHDTAVLLTRQD
 APRRCCHPRTPPCCGTSWLSRRRSMCI-----LAIPLPKCRGDYVGLGSRVPTPWCMTGSTRSLMLCVFRSD
 290 300 310 320 330 340
 400 410 420 430 440 450 460
 ICRADHKCDTLGLAELGTCDPRYSCSISBDSGLSTAFITANHELG---HVFMMRHDDNKKCEEGYKSPH
 --SPHASVLRGHNHQTQST--DSKRQLSTWIGTFELIGTCAVAHAGAVSNHFLNGFH-LTRCSHFVYALDH
 350 360 370 380 390 400 410
 470 480 490 500 510 520 530
 VMAPLNFYTPMMSKCSRRKYTFEFLDTGYECLLNEP-----SRPYRPLVQDLPGLIYVWVKQCELEIF
 KRONTSKGOYSMORRYVDKHCYTAHNDREVLYOEBSSNLETVHRWTADRP---PV-LFLCL-HHTAXCLF--
 420 430 440 450 460 470
 530 540 550 560 570 580
 GPGSVCPEYMOCR--RIMCMNV-NGVHKCGTQH-----TPWMDG-----TEDEGKIKCYKFGCVPEMD
 --PCHKPTEDKGIHRLFLAHQVLFSPHSRLRCIHFHPDIQONAAIGRQHRVYMMLEPRGQFSTGF-----
 480 490 500 510 520 530 540
 590 600 610 620 630 640 650
 VPVYDGSKGSWSPRTGCSRTGCGGKIVAIRBCNRPBRKNGSKYCVGSRMKFKCSQNTERRCLKOKRDRDEQCA
 HRAHCGGWMRS--GDSHPHGRG-----AELSHCLSGFEGMVTWNSPQOFLGYN--PGHHTSDRSR-----
 550 560 570 580 590 600
 660 670 680 690 700 710 720
 HFDGKHENINGL-LPNRYWPKYSGILMKDRCKLFCFVAGNATAYQLRDHYLDG-FPGGQDINDICVOGICR
 GHCSXHTSLAGLPIALHRSQSSS-----SPRPSIANHVSQNIILPLPL--VXAGHHRHSRQPSHLSGSXR
 610 620 630 640 650 660
 730 740 750 760 770 780 790
 QAGGDHVLNSKARADKCGVCGGDNSSCKYVAGTFNVLHYGNYVVRIPAGATMIDVRQHSFSGETDDNYLA
 ISSGYLSHKWLSMGSKTQIATGFLVHGMCFSSPFGQLEDDMMVTHSP---GYOLVXLLHYSHCHPLRY--
 670 680 690 700 710 720 730
 800 810 820 830 840 850 860
 LSSKGEFLNGNVVTMAKRETRIGNA-VEEYSGSETAVERINSTDRLEQEL--LLQ-VLSVSKLY-----
 KSOXKISILLAGIPVNHXNCTVPRDRPARMHIRPADKSAXMLMSQGHFNYLIDIHOTDTRPGGVALEFHR
 740 750 760 770 780 790 800
 870 880 890 900 910 920
 NPDRYSEFNLPTEDEK--PQDFYWNSHRQWQACSKP-----CQGERKRLKLVSTRESQDLTVSDGRCLRLP--Q
 XPSYLTLSIQSTVDKGTNCTSEMXLSLDVDCICCPVNHFGYXHKXFPSSSHHSFLINDIQSRONSASHPCYI
 810 820 830 840 850 860 870
 930 940 950 960 970 980 990
 PGIITTEPC--GTDCDLWVHVAASRESCAOCGLGYRRLDITCAKAYSRLDQGTTEKLVYDDGFCSSNHRKPSNRKRC-
 PLSIFHGCLWWDGCGCKNHQPSQSSHPACYIIMHSR--CPWCGMPSHTGHYIHSQWQAHATSGHSLYVRVQXCV
 880 890 900 910 920 930 940
 1000 1010 1020 1030 1040 1050
 ---SGECNTEGMYRWSAWTESKSCSDGGSTORRRAICVNTFRNDVLDSDSKT-----NOEKVTTQR-CSEBP-
 QAGAADRITFDQKQXADQIPWCKQOVFVSVPLGRVYC--MLMAHDCSSKTALEVYLDLEYYKKNVSRCCSTPFP
 950 960 970 980 990 1000 1010

1060 1070 1080 1090 1100 1110 1120
 -CPWKRGDWSECLVTCGCKGKHKRQVWCFCGEDRLNRCMDPRTKPKPMSQCOOPKPCASWQAGWGGCS--
 TKPRAKYLACDDL-----SKFFLRQ-----SRTHCTLPQ--HSCGEPFVW---PLXOOSFHL
 1020 1030 1040 1050 1060
 -----VTCGGQGYQR--AVKCIIGIYMSVVDNDNC--NAATRPDITDQCELPSCHPPAAEPTRRS
 ARHLVCLIKLSSGHRLEPLRNCALAAQWYMHOLEFGRPOYCNHNVLYXMFLEPCMKNYRHRHREHLTSLGL
 1070 1080 1090 1100 1110 1120 1130
 1180 1190 1200 1210 1220 1230 1240
 TYSAPRTQWRFGSWTPCSATCGKGTRRRYYVSC---RDENGSVADE--SACATLPRVAKRECSVYPCGOWKA
 SLKRDRIQ-----LAGKRGRRHYHLCSPGKKEFHSLCLEADSRCLSP-----LCRTTCG-----
 1140 1150 1160 1170 1180
 1250 1260 1270 1280 1290 1300 1310
 LDMSSCVTCGQG-----RATROYVCVNSDHYVDRSECDQDYPETDQDCSMSPCQRPDPDGLAQH----
 --PSS-FEHCIYXGSAHMEADR--XCXNAS-----RQSE-HTVHLGSLASAXDMAPC-CRTYIISFYVLAHSIFH
 1190 1200 1210 1220 1230 1240
 -----PQNEHYRPRRSASPSRTH-VLGGNQMRTGPGACSSTCGAGSSQRVYVVC
 HFLVLCSTLEWLEXPCLHMFMSRFQDSDNFR--IHLRSGHPLFMEKHIF-----CSAFQARPPSRHRPVC
 1250 1260 1270 1280 1290 1300 1310
 1360 1370 1380 1390 1400 1410 1420
 ODEGNYANDCVLEIKRDEORACSSGCPROMAYGNWGEETKLCGGGIRTRPLVVCQRNSGERPDLSCETLDK
 -AESG---SLCV-----LHKRYCSTRAVC--TASYAGTPEENLYOKSINHLYFSRCKGLAVGQAKGRWVIQVRYA
 1320 1330 1340 1350 1360 1370
 1430 1440 1450 1460 1470 1480 1490
 PRDEQC-NTHACNDAAWSTGTPWSSGSCGRGNKORNYCMKDKGSHLESDY---CKHLD-KPHGRKCR
 NTRNOCILKROXYLFDYTLTT-----SRGWCSSVLEEFYNAHG--SKLLDLTYCCCHHEACXTHG---PA
 1380 1390 1400 1410 1420 1430
 1500 1510 1520 1530 1540 1550
 GGRPKKAKAGWSSCSYSCGRCGVQQRHYGCOIGTHNKITAREPCEPRTYRPF-----SERDC-----Q
 HGRXKXKLYSIHYLHXHNSFYRDHKMPVQPLRYH-ICRELCAKRYVLLREOQYHDDGPHLLDCSSNAATGKSFLL
 1440 1450 1460 1470 1480 1490 1500
 1560 1570 1580 1590 1600 1610 1620
 GPRCLYTWRA-----EMWQ-----ECTKTCGEGSRYR-KVVCVDDNKNEVNGARCDVSKRVRDRESCGLQPC
 MLESHXKIKWNAHVHYESOLSSQXQYXLLNFQYLGILYRXRLDQLTLKLYSVGFLHGMKRFCCLPRLD--
 1510 1520 1530 1540 1550 1560 1570
 EYVWITGEMSEVPSWEL X
 TVDRIKKVVLSFCGSSSFLPCSPSYUHKOMKIPLLSRCLMLLNRDQXSFLLPFLMLFSSCLYCAVYCVLJRMH
 1580 1590 X 1600 1610 1620 1630 1640
 HAFLLRLSGRCAYKKKCGVCFVPLDHLVYRXYVVVQXNPNHETGMCRAFLSNGXKXKPPCAR
 1650 1660 1670 1680 1690 1700 1710

4. US-09-972-467-2 (1-1629)
 AB037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS: AB037733 5139 bp mRNA linear PRI 14-MAR-2000
 DEFINITION: Homo sapiens mRNA for KIAA1312 protein, partial cds.
 ACCESSION: AB037733

VERSION AB037733.1 GI:7242978
 KEYWORDS Homo sapiens brain cDNA to mRNA, clone_11b:pb1uescriptII SK plus
 SOURCE clone:fh11767.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (sites)
 Nagase,T., Kikuno,R., Ishikawa,K.I., Hirosewa,M. and Ohara,O.
 Prediction of the coding sequences of unidentified human genes.
 XVI. The complete sequences of 150 new cDNA clones from brain which
 code for large proteins in vitro
 DNA Res. 7 (1), 65-73 (2000)
 JOURNAL
 MEDLINE 20181126
 REFERENCE 2 (bases 1 to 5139)
 Ohara,O., Nagase,T., and Kikuno,R.
 Direct Submission
 Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba
 292-0812, Japan (E-mail:cdna@info.kazusa.or.jp,
 URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
 Fax:+81-438-52-3914)
 FEATURES
 source
 1. 5139
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="fh11767"
 /issue_type="brain"
 /clone_11b="pb1uescriptII SK plus"
 1. 4417
 /gene="KIAA1312"
 <1. 4417
 /note="Start codon is not identified."
 /gene="KIAA1312"
 /codon_start=2
 /product="KIAA1312 protein"
 /protein_id="BAA92550.1"
 /db_xref="GI:7242979"
 /translation="NSEHTAVLISLCSGMLGTFPSHDGDFTEPLQSMDDQDEDEEENK
 /PILYRRSAPQREPSSTGRNACDTESEKNNRHSKDKTRAKKWERLNLKADVYALNSG
 LATEARSAYGNKTDNTRERKTRHRTKRFSLYRPEVLVAVADRMYSYHGENLQHYIL
 TMSIVASLYKDPISIGNLIVIVLVIVLIVHNEEDGFSISFNAQTTLKRFQOMQHSKNS
 PGIHNDTVALLRDIDICRAHDKDPLGLALELGTIDDPYKSCISISDSGLSTAFITAH
 ELGHVNMHPDDNNKCKEKGYPQVMVPTLFLFYTNPMWMSKSKRYITFEFLDITGCG
 ECLINPERSRPLRVDQLGSLILYNNKQCELIIFGSCSQCPRYMKOCGRRLKMNVMNVGH
 KGRQTHTEPWADSTECPEGRNCKYGFVPEMDYRYLTDGSMGMSRFGTCSRPGCGGI
 KTAIRRCNRPREPNGKCYCVGRMKRKSCTEPCLEKQKRDREDCQANDNDGKHFNING
 LLEPNVMVPRKYSGLMKKDRCKLECRVAGNTAVYQLDRVDLTPGCGQDNLDLVOGLC
 ROAGCDHVLNSKARBDKCGVGGDSSCKTVAQTFMNTVHYGNTVVRIPAGATNIDVR
 OHSFSEPTDDNVALSSKGEFLLNGNFPVIMAKREIRIGNAVVERYSGETAVERIN
 STDRIQELLIDVLSGKLYNPRVRSFNPIEDKROQRPYHNSGPMQMSKRCQGER
 KRLVCTRBSDDQLTVSDQRDLRYPQGHITTEPGCTDCLMKNVHVASBESAOGLGTR
 TLDIYCAKYSRLDGKTEKYDDGFCSSHHPKSNNEKSGSEGTNGMYSAMTEKSCSD
 GGTORRRALCVNTRNDVLDKCTHOEKYTIORCSFEPGOMKSGMDSMLTYVCGGHL
 KHRQVWCSGFEEDRLNDRMDPRTKPTSMOICOOPEBCASMQAGMGGCSYTCGGQYDR
 AYKCIIGTYMSVVDNDCAARPTDPTQOPEPCSPKPPRAAPRTRETSYAPRTQYR
 GSWTPCSATCGKATRMRYVSCRDENKSVADSEKATLPRVAKRECSVYPCGOWKALD
 WSSCVTCGQGRATKQVWCVNSDHYIDRSECDQDYPETDQDCSMSPCQRPDPDGLAQH
 AQHPRNDEYRPRASASRTHVLYGQMRTEGPKLSSGSCAGSSQHRVYVCOENQYTL
 ANDCVBERIKRDEORACSESPQWALYGNMEECYKILGCGSITRFLVLRVANSSEPFDL
 SCEIILKRPDRQCNHACRPHDAWSTGPMWSSCSYSCGRCGVQQRHYVYVGOIGTHTKARETE
 DYCKHLAKRNGHNRKCGRGRPKKAAKANSQCSYSCGRCGVQQRHYVYVGOIGTHTKARETE
 CNPRTPESEEDCGPRCDYLTWRAEWEQECITKTCGEGSRYRKYVVCVDDNKNEVHGAR
 CDVSKRVRDRESCGLQPCSEYVWITGEMSEVPSWEL"

BASE COUNT 1471 a 1112 c 1362 g 1194 t
 ORIGIN
 AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803
 Initial Score = 8 Optimized Score = 80 Significance = 0.41
 Residue Identity = 108 Matches = 112 Mismatches = 333
 Gaps = 107 Conservative Substitutions = 3
 Translation Frame=

```

TVQVPSIILLIFPSYQDRI SAELTITNVIPIKMLNKEPVIPIEAVLLVIVDXVQLRSPMSWAKCILTLMMT
220 230 240 250 260 270 280
TTNKKKELRVPSSMLQHXITSTPTPCGQGVVVENISLSEPTLIVMASVCLLNINPDPDLLCLSNQASFTLXI
290 300 310 320 330 340 350 360
NNWKKFLDQVLRCHIXGSDGSGAITSMETTKAAGLSTHPGPKRSASLESTASMDVFPKPKMSEFXQMDP
370 380 390 400 410 420 430
GEVGVPLPEAPEHEVAGASKQPFESATDQONOKMVENIVXDXEYNLSPATRSVSSRSEPTSMNSVLLITGSL
440 450 460 470 480 490 500
TSTVCFRPGAGSLTIVTVEFXXRTGASCSAEMOGTPTISFETKXMMELIVARTOMISVSRFAAGKLDAIMEXT
510 520 530 540 550 560 570
OKPBEINVFVAIILHAKQOEHLIQYIMVTLIMSEFQVLPLIMCGSTVSOGKQDMMTTLXLYQAVKVNCS
580 590 600 610 620 630 640
XMFLLLSQWPKGRKALGMLKXSYVGRPLPKKELTQOIALSKNFCRERESCTTPMVAIILSTFQIKINLSS
650 660 670 680 690 700 710 720
FTGTVMGHKHAIVNPAKNGNENLIFABGNLISLLEFLIKDAIGCPSLDILLNIPVYQVYTXGGMIPAGVNVPS
730 740 750 760 770 780 790
VAWVTAHMTSTVPIAGMGRILRRLMNVFPAALIPNQTAVKNAQGNTRVAGAILPGLANQKAVTVGRREGEL
800 810 820 830 840 850 860
FVSIPEMMYMTANAHIKRRLPRFGAVSLVHSGNLETTGOSAMPVEKIGISTARSGVSLVIXIMIECVTLR
870 880 890 900 910 920 930
PSOHLCLVSRNVHPRGRVYVGDVAVSLVDRDTSXEBXNASLGLICOMKMTMVMQOOLDILPRTVYVHHVVI
940 950 960 970 980 990 1000
LPDLPKRRGGAHVHQPSSDGLIGRPAQPLVGRKVPBGXDTSAEMRMALMLTFRVPLPCIDQMKRKNVLXHP
1010 1020 1030 1040 1050 1060 1070 1080
VGNGRPMTGALALXPGVGOGRPKXCVSTVTTXSTIGVSVTRITISQKLTFTVYCHHALKGRQVVAALSTPSS
1090 1100 1110 1120 1130 1140 1150
MRTIIVPAPAPAPMCSVTSSEGLAPEGHEVYVPLADPSSVLLVYEMKMDTPOTVYWRKXNLMKREPVNPA
1160 1170 1180 1190 1200 1210 1220
10 20 30 40 50 60
LFTLLVLDLAEVGSPPAAA-----AVRKRDRLHPRQ--VKLELLETSEYEIVS-----PIRVNALGE
11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000
VLSGLMATAGESALS CVVEAAXEODMMSVSGPTVNGFOIXAVKFLINL---PIVSSVTHMLVHTTLDHGVLLLGA
1230 1240 1250 1260 1270 1280 1290
PEPPIVNHFKKTRRSIN---SAPDPMPAFASSSSSTSSQAHRYLSAFQOQFLFNLITANAGFA----PLFT
70 80 90 100 110 120
RVLSLV-----VEGINNEMFTAMQKMEAIKXVITVST-----WLSHMGESAEEDDAPKALDLSVALCP
1300 1310 1320 1330 1340 1350
V----TLIGTGVNQTIFYSEBEALHKCFKGYVWVNSHETAVTSLCSGMLGTFRRHDDYDIEPLQSDMEQ
1360 1370 1380 1390 1400
VADAYSRGMAVAVSEHTKXPERPSATHTPDRSHNATAKAH-GVPS---TLGGOR--NKNKNAPRP-----
1410 1420 1430 1440 1450 1460 1470
EDDEBEQKHIITIRRSAPQREBSTGHNACDTSHEKRNHSDKDKTKTARWGERLN-----LAGDVAALNSG
200 210 220 230 240 250 260
-----AAKAPGTAMCVMMTTKTRCMGHA-VTXASGRWTVKAAVVCNPAVSMGSGQENQORVYRPGNCNHRQLSHG
1420 1430 1440 1450 1460 1470
LATEARSVAIGNKNDNTREKRT---HRRTKRF--LSYRFVAVLVVADNRWVSHYGENLQHYITL--TLMSI
270 280 290 300 310 320

```

```

LRYAEKXVEGXYQECGRLEATRICHCELCVFXOYLSEFP-KLEFLVCKAGENXEMVSKLQYIILXFTLMS-
1480 1490 1500 1510 1520 1530 1540
330 340 350 360 370 380 390
VASIYKDPISIGNLINIYVILWVLIYHNEODGPISTSRNAQPTTLKNFCQ-WQHSKNSGGLIHHDTVAILLTQ-DI
--YVHKEEITGAAN---SXDAVRCIKVYKRRKILXCOGYNERNQDLXQSTFCYKGG-----AESI,SIQWLP
1550 1560 1570 1580 1590 1600 1610
400 410 420 430 440 450 460
CRANHXCDPLGLAELGTDIDPYRSCSISDSGLSTAF-TIAHELGHV- NMFHDDNNCKKEBGYK--SPQHY
IIIVYVYRQYQTFEKKRCKXKSGLVCEISXLENGXNFMFXNIVYRQKQTPRPNQWCCACILNNSFRGRKFSIIE
1620 1630 1640 1650 1660 1670 1680
470 480 490 500 510 520 530
MAPTLNFTYNNPMWMSKCSKTYTTERLDITGEGCLINBRESRPRYPLVQDPLGILVWVKQCELIFRPGSQVCP
VSOVPEF-----HISSCAQCLSRPFL-----MVINQXXS
1690 1700 1710 X
540 550 560 570 580 590 600
YMQQCRRLMGNVWVNGVYHKGCRPQHTPRMADGTCECFERKCKYGFYCPKEMDVPYVLDGSMGSMSPGTCGRTCG
610 620 630 640 650 660 670
GGIKYALIRECNPRPEKNGKCYCVGRMRKFKSCNTEPCLKQKRHFRDEOCANHFQKHFNFINGILRNVWVWPKY
680 690 700 710 720 730 740
SGILMKDRCKLFCVAGNVAAYQLDNDVYIDGTPGQDINDICVQDLCQWQACGDVNLKAKARADKCGVGGDN
750 760 770 780 790 800 810 820
SSCKTVAGTFNIVHYGYNTVVRILPAGATNIDVROHSFSGEDDDNYLALSSSKGFFLLNGNFMVVTMAKREIR
830 840 850 860 870 880 890
IGNAVVEYSGSTFAVERINISTRIEQELLQVLSYKLYNDVYRSEFNIPIEDKRPQCFYWNMSHGPMQACSFP
900 910 920 930 940 950 960
CGGERKRLVCTRESDDQLTVSDQRCDRLRQREHITTEPCGTDCLLWVHVASRSECSAQCGLGYRLLDICYAKY
970 980 990 1000 1010 1020 1030
SRLDGKTEKVDGDFCSSHPRKPSNRREKCSGECNWTGWRYSAMTPEKSCSDGGTQRRRAICVTRNDVLDSDSKC
1040 1050 1060 1070 1080 1090 1100
THQEVYTIORCEPFCPOWKSQDMSSECLVYTGCKGHKHROVWCQEGEDRLNRMWCDDEPTEKPKPSMOTCOQPECA
1110 1120 1130 1140 1150 1160 1170 1180
SMQAGPWCQCVYTCGQGYQLRAVKCIIGTYMSVVDNDNCNATRPRTDQDCELDPSCHPFPAPAPETRRSIVSA
1190 1200 1210 1220 1230 1240 1250
PRTQMRPFGSWPFCSATCGKGRFMRYVSGRDENGSVADBSACATLPRPVAKKEGCVYTPCGQWKAALDMSGCVY
1260 1270 1280 1290 1300 1310 1320
CGQGATRTQVQWCVNYSDVHIDRSECDQDIYIEPTEDDQCSMSPCQRTPTDPSGLAQHFPQMEDYRPPRSASPSRTH
1330 1340 1350 1360 1370 1380 1390
VLGQWQMTGPRGACASSYCCAGSGQRVYVVCODENYITANDCVERTIKRPEQVACSGSPQWMAIYCWMEGCTKVL
1400 1410 1420 1430 1440 1450 1460
CGGRTIRLIVVQCRSNGERFPLDLCCELLDKRPPDRQCNTHACPHDAMWSTGPMWSSCSVSGRHKORNVYCM
1470 1480 1490 1500
AKDGSHLSDYCKHLAKPHGRKCKGRCRPMK

```

5. US-09-972-467-2 (1-1629)
 ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139

TOIG of: ab037733 check: 4803 from: 1 to: 5139
 LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
 DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
 ACCESSION AB037733
 VERSION AB037733.1 GI:7242978

KEYWORDS Homo sapiens brain cDNA to mRNA, clone_1lb:pBluescriptII SK plus

SOURCE clone:fh11767.

ORGANISM Homo sapiens

REFERENCE Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo:

AUTHORS 1 (sites), Kikuno, R., Ishikawa, K.I., Hirose, M. and Ohara, O.

TITLE Prediction of the coding sequences of unidentified human genes.

JOURNAL XXI. The complete sequences of 150 new cDNA clones from brain which

code for large proteins in vitro

MEDLINE DNA Res. 7 (1), 65-73 (2000)

REFERENCE 2 (bases 1 to 5139)

AUTHORS Ohara, O., Nagase, T. and Kikuno, R.

TITLE Direct Submission

JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,

292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Kisarazu, Chiba

URI:ftp://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,

Fax:+81-438-52-3914)

FEATURES Location/Qualifiers

source 1..5139

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="fh11767"

/tissue_type="brain"

/clone_1lb="pBluescriptII SK plus"

gene 1..4417

/gene="K1A11312"

<1..4417

/gene="K1A11312"

/note="Start codon is not identified."

/codon_start=2

/product="K1A11312 protein"

/db_xref="GI:7242979"

/translation="NSEHTAVISLCSGMLGTFERSHDGDFIEPLQSMDEDEEEDONK

PHILYRSAPQREPTSTGRNACDYSERKNRHSKDKTKRARKWGERINLADVAALNSG

LATEAFSAYGNKTDNTRERKTRRTRKFLSYPFVAVLVADNRVSYGEMLOHIL

TLMSIYASIKDPSIGINLIIVIVLHNEEDGDSISENAQTILKNCOMOHSKNS

PGIHDHPFAVLRDODICRAHDKDPTGLAELGTGDDPDRSCSISPSDLSAFTTAH

MOEVSATLTLTLVLRDLAEMGSDPAAAAVAKRDLRHPROKLLLETISEVIYSPTRVNAIDGEPFTNYPKRT
80 90 100 110 120 130 140
RREINSATDPWPAFAFASSSSSSSTSSQNHRYLSAFGOOFLFNLTANMAGFTAPLFTVYLLDSTPGVNOYKPYSEEB
150 160 170 180 190 200 210
AELKHCFTFYGYVNTNSENHTAVISLCSGMLGTFERSHDGDFIEPLQSMDEDEEEDONKPHIITYRSAPQRE
220 230 240 250 260 270 280
STGRNACDYSERKNRHSKDKTKRARKWGERINLADVAALNSGLATEAFSAVGNKTDNTRERKTRRTRKRF
290 300 310 320 330 340 350 360
LSTPFEVAVLVADNRVSYGEMLOHILTLTMSIYASIKDPSIGINLIIVLHNEEDGDSISEPSISFPA
370 380 390 400 410 420 430
QTLTKNFCOMOHSKNSPGGIHHDVAVLRDODICRAHDKDPTGLAELGTGDDPDRSCSISPSDLSAFTTA
440 450 460 470 480 490 500
AHELGHVFNMRPHDDNNKCKEDEVKSPQHYMAFTLNFYTPMWMMSKCSKRYITIEFDLDTGEGCLLNEPESRPY
510 520 530 540 550 560 570
PLVYQLPGLILYVNNKQCELIIFGSGSQVCPYMMQCRRLMCMNNVNGVHKGGCRTOHTPWADGTECEPEKHKCYGF
580 590 600 610 620 630 640
CVREKMDVPTWDSGMSWSPRGTCISRTCCGGGKITKTAIRECNRREPANGKYCVGRMRMKFKSCNTEPCLKQKR
650 660 670 680 690 700 710
FRDEQCAHDEGKAFENINGLIPVNRVWPKYSGLIMDKRCKLFRVAGN-TAYVQLADRYIDGTGRCODINDIC
720 730 740 750 760 770 780
F-----FLXSCLL-----NRHNTSGSKYKGFVCPDLOCFTKNEFPHQVYKSHI--LVQIFSLSXKI
40 50 60 70 80 90
VQGLCR-----OAG-CDHYVNSK-----ARRDKCVGCGDNSCKTVAGT-----FNVYHYGY
790 800 810 820 830 840 850
RSGNAPRHFPCMATAGVKEILILPSSHCKXEVTVVAISPCCIPGTFITFSYTLXSLSSGTSQHPMSLQPLXXFLLYAN
100 110 120 130 140 150 160
NYVVRIPDAGATNIDVYQHSEFSGTDDNDYALSSSKGDFLLNMGNFVYMAKKEIRIAGNAVYVSSSEFAVER
170 180 190 200 210 220 230
RTSRXLTTECHREPESTXKPSLSTLPLCTIOGILAWMS-----XGLVRKHTVHSGK-----SGGXHQVSH
860 870 880 890 900 910 920
INSTDRLEDELLQVLSVSKLYIPVDVRYSEFNPIEDKPOQFTYNSHGR-WQACSAPC--QGER--KRLLYCT
930 940 950 960 970 980 990
IPALIPPLIPLP-----SGHGKXAD--GYS-----SODGTSDSHPVIQYVSOGCKLQLSRSTGRLLTSSO
1000 1010 1020 1030 1040 1050 1060
R---ESDQL--TVSDOR-CDRLPQPHGHTITPCGTCDDLRWN---VASRECSAQCGLGYRTLDI-----
1070 1080 1090 1100
RACPTSLFSLSTHTTTRKYLEPSPQVVLVHSCHSALQVXGHGHPQWSRSDSLGLVYGLHLSVSLAIIICLVXPXP
1110 1120 1130 1140 1150
YCAKYSRLDKTE--KVDDGFCSNHPKPSNRKESGECNTGGRKRSAMTEGSKSDGDTQRRAALCVNTR--
1160 1170 1180 1190 1200
TCECTPRPDQTHXIQADAFNHLGHP-----RHFLLCPGGLAR-CLQXSLSKWLPSPFAMQXTFRCLCPRPQ
1210 1220 1230 1240 1250
-----NDVLDSSKCTHDEKVTIQRC--EPPPCQWKSQDMSECLVLTGKHKHNRQVWCQEGEDRLADRMC
1260 1270 1280 1290 1300
ETHEHELQGPVLIHAASC-GQACVJHCPSRSGLSRIISQLKSGNRS-----PDRXQYTTSL
1310 1320 1330 1340 1350
420 430 440 450 460

BASE COUNT 1471 a 1112 c 1362 g 1194 t

ORIGIN

AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803 ..

Initial score - 8 Optimized score - 154 Significance - -0.41

Residue Identity - 19% Matches - 220 Mismatches - 668

Gaps - 249 Conservative Substitutions - 0

Translation Frame- 4


```

DEPTPTSMOTGQDREKASMOAGPWO--CSYTCGGQYQLRAVAKCITIGTWNYSVDDNDCAATPTD--(TOD
VLMPPHSLVHSPDLPYXAHXGQSPDQALCGSSGFLSTQSFSA-----VYFSSXHTTTRRMDPRAVQ
470 480 490 500 510 520
1160 1170 1180 1190 1200 1210 1220
CE-----LPSCHPRAPARTRSTYSAPRTQWRFGS-WTPGSAATGKG-----TMRRYVSCDENGSVADESA
LEHAPQGVLFHWF--PSTWVRLGLALR--GRXSSFMKGCXAKPLSGVLXGHDMEOSWSVSGIXSWSHSLR
530 540 550 560 570 580 590
1230 1240 1250 1260 1270 1280
CATLRRPAVAKERCSY-TPCGQWKAALDMSSCSVTGGGQ-----RAT--ROYMCVNYSDHVIDRSECDDYTP
STTSLALHTTTCVVALPDPQVTEQELSKAFHCPQGVTEHSSFPATGLGRAVAQALSSATPEFPSSRQLTYLT-
600 610 620 630 640 650 660
1290 1300 1310 1320 1330 1340 1350
EIDQDCSMPGCRGRTPDGSLAQNHPRONEDYRPRASPSRTHTVLGDNQWRGPRKMACSSTCAGGSGQRVYVCG
-----VYPRPQ-----YABGVQ--DPRNHVYLGAALYVLLVSGAAGGXHDGNSGWSVGLVALAQ
670 680 690 700 710 720
1360 1370 1380 1390 1400 1410
-----DENGATANDCVBRIRKPRDQ--RACESGPCPQWAXYMWGECETKRLGGGIRTRLY-----VQBRNSG
SLSSTPDLXVMMHFTALMSWYRCPQVYTHCGQACQDA--HSGCXQVC--IDVGLVSGSHLTSFNRSPP
730 740 750 760 770 780 790
1420 1430 1440 1450 1460 1470
ERFPLDSC-----EILDKRPD--REOCNTHACRPHDAAMSTGTPMSSCSVSGRGHKORHNVYCMAMDGSHL
NXHQWRKLCFPOVYTKHSDQSPRFHCGGQNSLHLMXWVTSXCVHLLSSSTMS-FRVLTO-IALLLWVPPQL
800 810 820 830 840 850 860
1480 1490 1500 1510 1520 1530 1540
-----ESDCKHLANR-HGHRKCRGRCRCPKWKAGAMSOCSVSGRQVQORHNGCOIGNHKIKARTECHNRY
FEHSAQAEKXRPQVPLVHSEHRSRLGLGWLKQSSFSFVPSLLYLAQXMSWVRKPKR-VALHSLLAT
870 880 890 900 910 920 930
1550 1560 1570 1580 1590 1600 1610
TRPSEKRCQG-PRCPRLYTWRAEMOBQCTKTCGEGSR-YRKKVVCVDDNKNKEVHGARDVSKRPVDRRECSLQ
CHLRQSQVYPOGSMVACP-----GWSRSRHLXSEFVXSQSDSLVQTSFPRFRSPMOGLLHAGHGP
940 950 960 970 980
1620
PCFYWITTEGSEVPSWEL X
990 1000 1010 1020 1030 1040 1050
XLFQXNCXGLSSIGILEKEXRTSGLYNEPPTDKTXSKSSCSMRSEVELLSTFVSDPIYSTTAFPMRISLALIV
1060 1070 1080 1090 1100 1110 1120 1130
TKRPFKNSPLLDLAKKXSSSSVPEKLCRTSILVAPAGIRTTVLXPKCTLVLVNVPATVYHHEELSPQPTPHL
SLRAFEKTXSHRACRORPWTQISFVSWPQGVPSITLSRSKXXXAVFPATLQNNMLHRSFIRIPLYLGMORTLS
1140 1150 1160 1170 1180 1190 1200
SRPLMLKCFPSKXAHGSSRKRSPCLRHGSLVLDLNTILRPTQYFPPFPFGSGLLHSMRMAVLMPPHVLEQVPK
1210 1220 1230 1240 1250 1260 1270
GLDLPDQPSVYTGNSISLGTQNPYLQCFPPSGSHSVPSAQGVXVRPQPCPTPLTLLHOSRIMCITTYGHXHEGPK
1280 1290 1300 1310 1320 1330 1340
1350 1360 1370 1380 1390 1400 1410
INSKCLFTLXRMGSGWTGKXGGLDSSGLSKHSPXRVSKNSVLYFRLHFHDIQGLVXKFSVAGMTCCGILLPS
SLHLLSSXGMLWTPSSMAIVKAVLNLPLSSLLEOLLXGSOAWPSSARPKPVSHLSXALDIOICLIVKRTAVYSXW
1420 1430 1440 1450 1460 1470 1480 1490
IPRGLFEECCWQKFRVYVYALXADMGGPSCSLXITTKFTTTLIKRPLIGLSLXIEATFIDIKYKIXCXRFSPW
1500 1510 1520 1530 1540 1550

```

```

YFTLLSATTKTSMNRGXDKRRFVLLWLFSLVLSYLRPAENASVAKPLFNAAATSPARLLISPHFLALVF
1570 1580 1590 1600 1610 1620 1630
LSLCLFLESEVSHACLPVHSGSLMGALRLXMMWGLFCSSSSGSSIDCSGSIKXSPDXDRNVPISPEQRLMT
1640 1650 1660 1670 1680 1690 1700
AVCSSE
1710

```

6. US-09-972-467-2 (1-1629)
 ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
 TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
 DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
 ACCESSION AB037733
 VERSION AB037733.1 GI:7242978
 KEYWORDS
 SOURCE Homo sapiens brain cDNA to mRNA, clone_lib:pbLuescriptII SK plus
 clone:fh11767.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (sites)
 Nagase,T., Kikuno,R., Ishikawa,K.I., Hitosawa,M. and Ohara,O.
 Prediction of the coding sequences of unidentified human genes.
 XVI. The complete sequences of 150 new cDNA clones from brain which
 code for large proteins in vitro
 DNA Res. 7 (1), 65-73 (2000)
 2 (bases 1 to 5139)
 Ohara,O., Nagase,T. and Kikuno,R.
 Direct Submission
 Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba
 292-0812, Japan (E-mail:cdna@foc.kazusa.or.jp,
 URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
 Fax:+81-438-52-3914)

FEATURES
 source
 gene
 CDS

Location/Qualifiers
 1..5139
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="fh11767"
 /tissue_type="brain"
 /clone_lib="pLuescriptII SK plus"
 1..4417
 /gene="KIAA1312"
 <1..4417
 /gene="KIAA1312"
 /note="start codon is not identified."
 /codon_start=2
 /product="KIAA1312 protein"
 /protein_id="BA092550.1"
 /db_xref="GI:7242979"
 /translation="NSEHTAVISLCSGMLTGFPSHDQDYFIEPLQSGMDEDEDEEQRQNK
 PHILYRSAQDREPSTGRHACDTESEHNRHSRKRKTRAKRGERINLAGDVALNSG
 LAPEAFSAGKDPNTRERKTRHRTKRELFYPRVEVLYAADRMYSYGEMNHQYTL
 TMSIVASITKDDPSIGMLTIVYVNLIVTINDEGDSISFNQOTYTKRNCQWHSKNS
 PGIHDTAVLLRQDTCRAHDKCDTIGLELDTICDPYKSCSISDSGLSTAFTHA
 ELGHVEMPHDDNNKCKEVEGKSPQAHMAYTLNFTYNNPMMWSKCSKRYITFEFLDTYG
 ECLINDEPSPYLPVQDLPGLILYVNVKOCFLIDPGSOVCPLYMOCRLMKNVNVGH
 KGGTQHTPMAADSTECRPNKCYGFCVPEKEMDVPYLDGSDMSGMSDFGKCSRGGGI
 KTAIRECNRPEKNGGKYCVGRMFKRSCQTEPCIKOKRDPDRQCAHHRDCKHFNINQ
 LPAKRVNPRKSGILMKDRCKLTCRAAGNATAYQLDRAVITDGTGCGDINDICVQGLC
 RQAGCDVLSKARRDKCGVGGDSSCKTVAQTEFTVHGYVTVVIRIAGAVNIDVR
 QHSSEGTDDNDVALSSSKGEFLLMGNFVYMAKREIRIGNAVVYSSSEFAVERIN
 STDRIBEQLLQVLSVQKILNPDVRYSPFNIPIDKQQLQYVMSHGQWQACSKRPNVGRIN
 KRKLVCTRESDDQVLPDQCDRLRQPGHLEPCGTDCLRMHVAASRSGSAAQCGGER
 TLDIYCAKYSRLDQKTEKVDGFCSSHSPKRSNERKSGCQNTGGMGYMAYWATGCSKSD
 GGTGRRRALGVNVRNDVLDSSCKTHQEKYTIQRCSFEPQPKMSGDMSECLVTCGKGH

KIRQVWCQGEEDRLNDRKMDDETRKPTSMOTQOQREPCASQWAGPMSQCSVTGGGQYLR
 AVKCIIGITMSVVDNDNNAATRPFDODCELPDPPAPAEKRSYTPAQRQMR
 GSWTPCSATCGKGTMMRYVSCRDENGSAVDEBACVTLDRPVAEKRSYTPGGMWALD
 MSSCGVCGQGRATQAVK VNYSDHVIDRSECDDYIETDODCSMSRCPDRTDGL
 AQHPONEDYRBRASPSRTIHLGDNOMRTGMSACSSSTAGSGQRVYVCGODENGYI
 ANDVCEKRPDORACESGPGQOMAYGNMGBEOTKICGGGIRTRIVLWCCORNSGEBPL
 SEEILDKPDRDRCQCTHACRPHDAAMSTGPMSSQCSVSGKRVQOQHNGCQIGTKIARETE
 DYCKHLAKPHGKRCGRCPKAKAGAWSQCSVSGKRVQOQHNGCQIGTKIARETE
 CNPYYRPESEDDCGRCPCELYTRAEEMQECTKCTGEGSSRYRKYVAVDNNKNEVHGAR
 CVSKRPVYRESCSIOPCSEYLVITGEMSWPSPWML"

BASE COUNT 1471 a 1112 c 1362 g 1194 t
 ORIGIN

AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803

Initial Score = 7 Optimized score = 205 Significance = -0.41
 Residue Identity = 188 Matches = 303 Mismatches = 999
 Gaps = 325 Conservative Substitutions = 0
 Translation Frame= 6

MQFVSMATLLTLVYRDLAEMGSPDAALAAVRRKDRLHRQVKLLETLSKEYEIVSPDIRVVALGEPRTVWHFKRT

RRTNSATDPPAPASSSSSTSQAHYRLSARSGQDFLNLGNAGFIADLTVTLLGTGVQVQTKFYSEE

AELKHCFYGVYNTNSEHTAVISLCSGMLGTRFSHDGDFY---EPLQSMDEQED-----EEESON

DYWFITIKKGLDKKCAQEDIKKNGTXXDPSDIENYF

KPHIYRSARQREPRSTGRNACDTSYSEKHNRSKDKKTKRARKWGERINLADGVAALMSGLATFAFSYAGKT

LPLKLFVQAQHIWPMKGVCLPLT-----MFXNQKFOPEPSSXEISHTSPDFXPL-----LKNQVWX

DNTREKTRHRRTRKRFSLYRFEVLYV-VADNRMVSYGEMLNQHYILT-LMSIVAS-TYKDP-SIGMLINLVI

RKTTEFLHSHMSER-DSAPRFQHVLYVCHSCHLSLILYPMHNYIIFLYLIVFRHLTKASYEFAPAVIISISLQXDI

VNLIVIHNEODQSPIS--FNAQTT-----LKNFCOMQHSKNSPGGINHDTAVILLTRQDICRAHDKCDTL--

KVNHRMXXRDLLETISYISPALHTRNSSLSGLKLRCCOKTHS--SOMOI-GRVASSLPRPSCTHNPTNPSATLPR

----GLAELGTICDPRYRSCSISDPSGL-STAFTIATHELGHVFNMP-HDDNNKCKEKGVSPOHVA---PT

WLSXKMYLQRPKRYLXP-FSCDRDLIAGIQTTAFTVNRHLAHTACRPHILVYVHTHHLAVRGAFAAGLDAF

LNFYINPMMWSKCSKRYITTEFLDITGYGECILLNEPESRYPRLVQDPLGLLVNWKQCEL-----TEGPGSOV

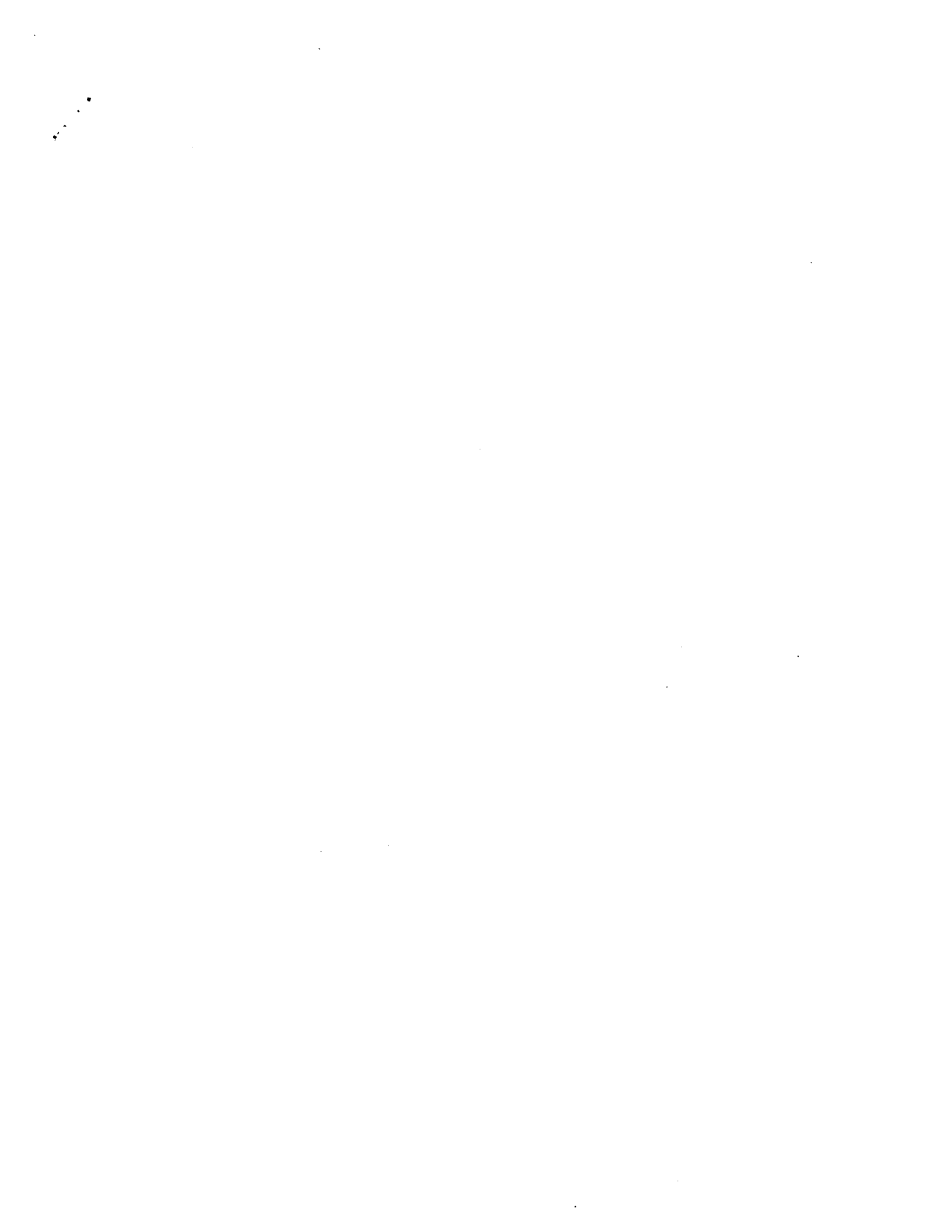
LPFLCPRVEGPMALAVAFRLRSGVWVALGLSGYFVCSDLTAMHPLLAYSATGHRALITPSASPPGASS--

--SSALSTVPMMLSOVLVITFTXMASIFCHAVNISLFPSTTRDRTRAP---RASTFCSVVTSMCVT-LIT

CGGGIKR--TAIRECNRPRPKNGKGYCVGRMMKFKSCNTEPCLKQKRDPRDQGANHPRDKHFNINGLLPRNWK

IGRFKKNTTA---QIMKRFVGLPLTDHQSCSYASTTQLSALSPVALISPLRTRRAGFTGSLLDIRPYSHHTV--

VPKYS-GILMDKRCLEFRVAG-----NFAAYQLRDRYIDGTPEGGQDITDI-----CYQGLCRQAG
 680 690 700 710 720
 YGVYIFILITVNNPDLGASVYGTGTCSPGASSPLVSTEGHMAAGAGAGBTLLVLEGLVSKATVWGLRBMX
 510 520 530 540 550 560 570
 CHVYLSKARRDKCGVCGGDNSSCKTAGTSTNTYHYGYNYVYRIPAG-ATNIDVROHSFSGEPTDDNYLALS
 730 740 750 760 770 780 790
 HGTVLVVS--FWMIIILVTLPLRPHVYVVDT--HNLPGCPY--LPTGRARAPAVQGLRPLPFGCHRTFPLCHW
 580 590 600 610 620 630 640
 SSKGEBLLGNVFTVMKRELRIGNAVYVYSG----SETA-VERINSTRIDRIEQLLDLVLSYGLKLYNDVRY
 800 810 820 830 840 850 860
 SHQG---SYGTLY---SHRALIISADVYSHDPTFTSGXAWGPRKRS-----PLGSWCTV--CASBFRFG
 650 660 670 680 690 700
 S--FNIPLEDKPQDFYVNS-----HGFWQA--CSKPGQGEKRRKLYCTRESDQLTVSDQKCDRLP
 870 880 890 900 910 920
 SWGRMTXMXKFTVLGISWSSCCITVIVYHXHISPNDAFRCS-----XVLSLSTSD---YALSPEGRIDP
 710 720 730 740 750 760
 OPHHTPEPCGTPDCDLRKHVYASRSECSAOCGLGRTLPD-ITYCAKYSRLDGTKEKYDDGFCSSHP-----KP
 930 940 950 960 970 980
 SREKCSGEGCNTGGRY--SAMTECSKSGDGS-----GQRARRACVMTBNDVLDSSCTHOEKYTIQRCS
 990 1000 1010 1020 1030 1040 1050
 LN-GNEFLMCAEPAVIOYIISIGIDTNSPSPGLGPTVAEYTFPESGRIDARPR-----VT--FPAK
 830 840 850 860 870 880
 FPCPGKSSGDNSECIVTGCCKGKHNQWQCGEGEDRLNBMQDPEPKRPMOTQOQREPCASWQ--AGPWGQCS
 1060 1070 1080 1090 1100 1110 1120
 FFTVAV--FGMAAAKTIINLLSLRPIPAIFG---TVDYQC-AVTQATILGTTFTAGGMMRPOQVYVCTGTFPSN
 880 890 900 910 920 930 940
 V-TCGQ--GYQLRAVKCI---IGTYKSVYVDDNDCAAPR-TRDQDCEL-----PSCHPRPAAPETRRSTYSA
 1130 1140 1150 1160 1170 1180
 VSRIGPRLASLRNSKLLIRFPQANKFSPPLAGFTACLIPRMYTVRKYLLRTEFMNIBERLAIYGVVDSHR
 950 960 970 980 990 1000 1010
 PRTQWRFGSMTPC---SATCGKGTBRMYVS-----CRDENGSVADDESACATLPRVAKEEGC-----SVT
 1190 1200 1210 1220 1230 1240
 QNLKQKFLNALICXVNSYFSGSLGPTVLYHSIRNANPFGHCNKKYSIXOEFPTGAXXSVVYVYGFYKRYTLV
 1020 1030 1040 1050 1060 1070 1080
 PCGQWKAALDMSSCSYVCGGGRATROVMCVNYSDBHYDRSECDQYIPEFDQCSMSP--CPRQRTDPSGL---
 1240 1250 1260 1270 1280 1290 1300
 PHINIGSTWNS-----DHSIVTITMYCIKCSCH-----CFAKXIATPTNPTFISPEFYXVNMILASSLPRK
 1090 1100 1110 1120 1130 1140
 AQHPRQNDYRPRASAPSRTHVILGQWQMTGPMGACSSNCSAGSQRVYVCGDDENGYTANDCVERIK--PDE
 1310 1320 1330 1340 1350 1360 1370
 ALDITICVLAIRS---SIHYSYKLIYGCVP---CHS---ABQLAPVLDH--NSTYFRDPAHDKQKQTVDY
 1150 1160 1170 1180 1190 1200
 QRACEGSPCQMAVYGMWGBCTKLCGGG---IRTRLVYVQCRSNGERPRDPSCEIILKRPDRQECNTHACRPHDA
 1380 1390 1400 1410 1420 1430 1440
 KMLRQVYVSTLFEIVSLLLETVLRAVAGLKHSTSTVYFSTITMFWMSVALSNGCFDAP-----STGCSG
 1210 1220 1230 1240 1250 1260 1270



> O <
O I 10 Inteligenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file 09972467-2-289-478-ab037733.res made by jdelaval on Thu Jun 10 6:16:17-

Query sequence being compared: US-09-972-467-2 (289-478)
Number of sequences searched: 6
Number of scores above cutoff: 6

Results of the initial comparison of US-09-972-467-2 (289-478) with:
File : ab037733.seq

Table with 10 columns: S, E, U, F, N, O, P, R, S, N, U, M, N, C, E, S. Rows include similarity matrix, mismatch penalty, gap size penalty, etc.

PARAMETERS

Similarity matrix: Unitary 2
Mismatch penalty: 1.00 Joining penalty: 20
Gap size penalty: 0.05 Window size: 32
Koff score: 0
Admization group: 0

SEARCH STATISTICS

Scores: Mean 37 Median 9 Standard Deviation 74.55
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00
Number of residues: 10271
Number of sequences searched: 6
Number of scores above cutoff: 6

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% similar sequence to the query sequence was found:

Init. Opt.

Table with 5 columns: Sequence Name, Description, Length, Score, Sig. Frame. Row 1: ab037733, TOIG of: ab037733, check: 480, 1712, 190, 190, 2.05, 2

The list of other best scores is:

Table with 5 columns: Sequence Name, Description, Length, Score, Sig. Frame. Rows 2-6: ab037733, TOIG of: ab037733, check: 480, 1712, 8, 30, -0.39, 4; ab037733, TOIG of: ab037733, check: 480, 1712, 8, 34, -0.39, 3; ab037733, TOIG of: ab037733, check: 480, 1713, 8, 28, -0.39, 1; ab037733, TOIG of: ab037733, check: 480, 1711, 7, 37, -0.40, 6; ab037733, TOIG of: ab037733, check: 480, 1711, 6, 33, -0.42, 5

1. US-09-972-467-2 (289-478)
ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139

TOIG of: ab037733 check: 4803 from: 1 to: 5139

LOCUS: AB037733 5139 bp mRNA linear PRI 14-MAR-2000
DEFINITION: Homo sapiens mRNA for KIAA1312 protein, partial cds.
ACCESSION: AB037733
VERSION: AB037733.1 GI:7242978
KEYWORDS: Homo sapiens brain cDNA to mRNA, clone_lib:pluascriptII SK plus

ORGANISM

Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE

Nagase, T., Kikuno R., Ishikawa, K.I., Hirosewa, M. and Ohara, O.
Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
DNA Res. 7 (1), 65-73 (2000)

JOURNAL

Ohara, O., Nagase, T. and Kikuno, R.
Direct Submission
Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba
URL: http://www.kazusa.or.jp/huge/, Tel: +81-438-52-3913,
Fax: +81-438-52-3914

FEATURES

Location/Qualifiers
1..5139
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="fhl1767"
/issue_type="brain"
/clone_lib="pluascriptII SK plus"
1..4417
/gene="KIAA1312"
<1..4417
/gene="KIAA1312"
/feature_start="KIAA1312"
/codon_start=2
/product="KIAA1312 protein"
/db_xref="GI:7242979"
/translaton="NSEHTAVISLCSGMLGTFPSHDDVYFIEPLQSMDEDEBEONK
PHITRRSAPORREPTSGRHRACDPTSEHKRRSKKTRAKRWGERINLADVAALSG
LATRARSAYGNKTDNTEKTRTRTRKRFLESYPRPEVLYVADNBMYSYGENIQHYIL
TMSIVASITKDPISGLNIVIVLVNLIHNEODGSPISFAQTTLNRCQWHSKNS
PGGIHDTAVILTRDRCRAHDCDITGLAELGTIDDPYRSCSISBDSGISTAFITAH
ELGVFPMHPDDNNKCKEVEGSPQHVMAPTLNFYINPMWMSKCSRSKYLIEFDGTYG
ECLINPESSRPIYPLVQDGLIYNVKNOCLELIFGPSOVCPYMMOCRRLLMCNNVNVGH
KCGTRQHTPADGTECPGRCRHKCKYGVCPREMDYVPTDGGSGMSRPGCSRPGGGI
KTRTRCNRPPEPKNGKRTCYGRMKRKSCTREPCLOKRPFRDEQCAHFDGKRFNNG
LIPVWVWPYISGLMKDRCLKFCRVAAGNTAVYQLDNRVIDGTGPGCGQDINDICVQGLC
FOAGCDHVLNSKARRDKCGVCGDNDSSCKTAVAGTFMVTYHYGVTMVAIPAGANNIDVR

Handwritten signature/initials and date: 6/20/02

7

OHSFSEEDDDNYLALSSKGEFLANGFVYMAKREIRIGNAVEYSSSETAVERIN
 STDRIEQLLELVYLVKLYNPDVRYSPNIPIEDKPOQRYNNSHPWQASPCGGER
 KRLLVCTRESDDLVSDORCDRLPQPHGHIITTECGCDLIRHVAIRSESSAQCQYR
 TLDIYCAKSRLLDGETEKYDDGSSHPKPSNRKSCGSCGNTGMSRYSMTSCSSCD
 GGTORRRRAICVMTNRNDVLDSSGCTHREKATYORCSERPPOKMSDMSBCLTCKGKH
 KHQVWVCGEGRDLNDKCAATRPPTKOTQPEASAOQAPWQCVSTYCGQYLR
 AYKCLIGITMAYVDDNDCAATRPPTODCELPSCHPPPAEPTRRSYAPRTQMR
 GSWTPCSATCGKGTMRVYVSCDENGSVADDESACATLPPVAKRECSVTPCGOMKALD
 WSSCVTCGGRATQKVCVNYSDHVIDSECDQYIPEPTDDGMSFSCPORTPISGL
 AOHPPFEDYRPRSPASPRTHVLDGQWMTGSMGACSSRTKCGGIGIRLVLCORSMERFDL
 ANDCVERIKPDRORACEGSPGPOAYVGNMSECTKICCGGIRLVLCORSMERFDL
 SCETLDKPPDRRCQNTHACRPHDANSTGFWSSGSCGGRHQRVNYCAMDGHLES
 DYCKHLAKPHKCKGRCRPMKAGAMQCSVSCGRVQORRVCQIGTHIARETE
 CNPRTPESEKDCQGRCPLYTMRABEMQECTKTCGEGSRVRYKVVCDNKNREYHGAR
 CYSKRPPVRESLSQPCRYVITGEMSEVSPMEI"

BASE COUNT 1471 a 1112 c 1362 g 1194 t
 ORIGIN
 AB037733 Length: 5139 June 20, 2002 06:02 Type: N Check: 4803
 Initial Score = 190 Optimized Score = 190 Significance = 2.05
 Residue Identity = 100% Matches = 190 Mismatches = 0
 Gaps = 0 Conservative Substitutions = 0
 Translation Frame = 2

NSEHTAVISLCSGLGTFRRSHDDGYTEPLQSDMEDEEONKPHIYRARSAPQREPTGRHACDTSKHN
 10 20 30 40 50 60 70 80
 RHKDKKTRARKKGERINLAGVVAALNSGLATEAFSAVGNKTDNTRERKTRRTRKRFIISYRFEVYLVAD
 80 90 100 110 120 130 140 150
 NRMVSYGENLQHYIITLMSIVASIKKPSIGNLIVYLVNLIYVHNEQDGSISNAOTTLKRFQOMQSK
 160 170 180 190 200 210 220 230 240 250 260 270 280
 NRSVSYHGEMLQXITLMSIVASIKKPSIGNLIVYLVNLIYVHNEQDGSISNAOTTLKRFQOMQSK
 150 160 170 180 190 200 210 220 230 240 250 260 270 280
 NSGCGIHHDTAVLLTRQDICRAHDKCDTGLGELGTLICDPYRSCSISSEDSGLSTAFTHAHELGHVFNMFHD
 90 100 110 120 130 140 150
 NSGCGIHHDTAVLLTRQDICRAHDKCDTGLGELGTLICDPYRSCSISSEDSGLSTAFTHAHELGHVFNMFHD
 290 300 310 320 330 340 350 360
 KOCELLIFGEGSOVCPYMOCCRMLMNNVNVHKGCRTOHTPWADGETECPGKHCYKGFVCPKEMDVPYTDGS
 370 380 390 400 410 420 430
 WGSMSPEGTCSRTCGGGITTAIRECNRREPRKNGKYCVGRMRKFKSCNTPEFLKQKRPDRDEQCAHEDGKHF
 440 450 460 470 480 490 500
 NINGLLPWRWPKYSIGILMKDRCKLFCFRVAGNTAYYOLRDRVINDGTPECGDPTNDICVQGLCRQAGCDHVLN
 510 520 530 540 550 560 570
 SKARPDKCGVCGDNDSSCKTVAGTEFNTVHYGNVYVRIIPAGATNIDVQHSSEGETDDNDLALSSSKGEFL
 580 590 600 610 620 630 640
 LINGNEVYVYAKKEIRIRIGNAVEYSSSETAVERINSTDRLEQELLQVLVSVGKLYNPDVRYSPNIPIEDKPOQ
 650 660 670 680 690 700 710 720
 FYWNSHGHPWQASCKRPOGERRKRLVCTRESDDLVSDORCDRLPQPHGHIITTECGCDLIRHVAIRSESSAQCQYR
 730 740 750 760 770 780 790

CGLGRVTLDIYCAKSRLLDGETEKYDDGSSHPKPSNRKSCGSCGNTGMSRYSMTSCSSCDGTDGQRRRA
 800 810 820 830 840 850 860

ICVNTRNVDLDDSKCTHOEKYTIORCSFPCPOKMSGMSBCLYTCGKHKHROWCOFGEDRLNDMCPDE
 870 880 890 900 910 920 930
 TPRTSMQTCQDEPCASWQAGPWCQSVTCGGQYOLRAVKCIIGTYMSYVDDNDCAATRPPTDQCELPSCG
 940 950 960 970 980 990 1000
 PPAAPERTRRSTYSAPRTQWPRGSGTPECSATCGKTRMRVYVSCDENGSVADESACATLPPVAKKECSVTP
 1010 1020 1030 1040 1050 1060 1070 1080
 GGQKALDMSSCSYCGGGRATROYMCVNSDHYIDRESCDQDYPITPDQDCMSPCPQRTPPDGLAQHPFQ
 1090 1100 1110 1120 1130 1140 1150
 NEDYRPRASPSRTHVLDGQWMTGSMGACSSITGAGSQRVRYVVCODENGTANDCYERIKPDRORACEGSP
 1160 1170 1180 1190 1200 1210 1220
 CPQWYGNWGECTKICGGGIRTRLVVCCORSNGERPPDLSCHILDKPPDRRCQNTHACRPHDANSTGFWSSCS
 1230 1240 1250 1260 1270 1280 1290
 VSCGRGHRORRNYVYCAMDGHLES
 1300 1310 1320

2. US-09-972-467-2 (289-478)
 ab037733 TOIG of: ab037733 check: 4803 from: 1 to: 5139
 TOIG of: ab037733 check: 4803 from: 1 to: 5139
 LOCUS AB037733 5139 bp mRNA linear PRI 14-MAR-2000
 DEFINITION Homo sapiens mRNA for KIAA1312 protein, partial cds.
 ACCESSION AB037733
 VERSION AB037733.1 GI:7242978
 KEYWORDS
 SOURCE Homo sapiens brain cDNA to mRNA, clone_lib:pluescriptII SK plus
 clone:fh11767.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (sites)
 Nagase,T., Kikuno,R., Ishikawa,K.I., Hirosewa,M. and Ohara,O.
 Prediction of the coding sequences of unidentified human genes.
 XVI. The complete sequences of 150 new cDNA clones from brain which
 code for large proteins in vitro
 DNA Res. 7 (1), 65-73 (2000).
 20181126
 2 (bases 1 to 5139)
 Nagase,T., Nagase,T. and Kikuno,R.
 Direct Submission
 Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
 URL:http://www.kazusa.or.jp/huge/, Tel: +81-438-52-3913,
 Fax: +81-438-52-3914)

FEATURES
 Source
 Location/Qualifiers
 1..5139
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="fh11767"
 /rname="fh11767"
 /rsize_type="brain"
 /clone_lib="pluescriptII SK plus"
 1..4417
 /gene="KIAA1312"
 -1..4417
 /gene="KIAA1312"
 /note="Start codon is not identified."
 /codon_start=2
 /product="KIAA1312 protein"
 /protein_id="BA02550.1"
 /db_xref="GI:7242979"
 /translation="NSEHTAVISLCSGLGTFRRSHDDGYTEPLQSDMEDEEONK
 PHIYRARSAPQREPTGRHACDTSKHNKPKKDKKTRARKKGERINLAGVVAALNS
 LATEAFSAVGNKTDNTRERKTRRTRKRFIISYRFEVYLVADNKNREYHGELQHYI
 TMSIVASIKKPSIGNLIVYLVNLIYVHNEQDGSISNAOTTLKRFQOMQSKNS
 FGIHHDTAVLLTRQDICRAHDKCDTGLGELGTLICDPYRSCSISSEDSGLSTAFTHA

CDS
 gene
 CDS

CGLGRVTLDIYCAKSRLLDGETEKYDDGSSHPKPSNRKSCGSCGNTGMSRYSMTSCSSCDGTDGQRRRA
 800 810 820 830 840 850 860

